

### **REMARKS**

Upon entry of this amendment, claims 1-29 are pending. Claims 30-33 are canceled. Claims 1, 13-24, and 26-29 have been amended. Applicants respectfully submit that the amendments do not introduce new matter and are made without any intention to abandon the subject matter as filed, but with the intention that claims of the same, greater, or lesser scope may be filed in a continuing application.

### **Rejections Under 35 U.S.C. §112, First Paragraph**

The Examiner rejected claims 19-21 under 35 U.S.C. §112, first paragraph, contending that they contain subject matter that was not described in the Specification in such a way as to reasonably convey to one skilled in the art that the inventors had possession of the claimed invention at the time the application was filed. Applicants traverse the rejection to the extent it is maintained over the claims as amended.

Applicants submit that their specification provides support for a component that associates with an RNP complex if it associates with a  $K_d$  of about  $10^{-6}$  to about  $10^{-9}$ , about  $10^{-7}$  to about  $10^{-9}$ , and about  $10^{-8}$  to about  $10^{-9}$  in paragraph 44 of the Specification. Applicants submit that they do not require drawings of such structures for binding and that the dissociation constant  $K_d$  of such interactions is well known in the art or is easily determined by a skilled artisan. For example, Applicants submit herewith a number of references that predate the filing date of the instant application that indicate that binding assays for determining  $K_d$  of RNA binding proteins were well known in the art and that a skilled artisan would be well aware of the  $K_d$  between an RNP complex and its components.

For example:

“*FMR1* Protein: Conserved RNP Family Domains and Selective RNA Binding” Ashley et al. (1993) Science 262:563-66: Columns 2 on page 565 describes the  $K_d$  for FMR1 mRNA of between about 5.7nM and 39 nM. (Exhibit A).

**“AUF1 Binding Affinity to A+U-rich Elements Correlates with Rapid mRNA Degradation”**

DeMaria et al. (1996) J. Biol. Chem. 271:12174-84: Columns 1 on page 12181 describes the K<sub>d</sub> for Rβ and (AUUU)<sub>5</sub> of about 210±50 nM to 19±7nM. (Exhibit B).

“The Neuronal RNA Binding Protein Nova-1 Recognizes Specific RNA Targets in Vitro and In Vivo”: Buckanovich et al. (1997) Mol. Cellul. Biol. 17:3194-3201: Column 1 of page 3196 describes the K<sub>d</sub> for NFP for SB2 to about 2 to about 20nM and the K<sub>d</sub> of N3 for SB2 is about 10 to about 180 nM. (Exhibit C).

“The Determinants of RNA-binding Specificity of the Heterogeneous Nuclear Ribonucleoprotein C Proteins” Görlach et al. (1994) J. Biol. Chem. 269:23074-78: This article describes a number of K<sub>d</sub> measurements, for example, on page 23076 column 2, the K<sub>d</sub> of U2AF for different 3' splice site regions is 10<sup>-8</sup> to 10<sup>-6</sup>M; on page 23077 the K<sub>d</sub> of hnRNPC1 for CB2 is 170 nM. (Exhibit D).

“Binding Protein: Localization, Abundance, and RNA-Binding Specificity” Görlach et al. (1994) Exp. Cell Res. 211:400-407: This abstract describes that hPABP binds to oligo(rA)-rich sequences with a K<sub>d</sub> of 7nM. (Exhibit E).

“Wheat Germ Poly(A) Binding Protein Enhances the Binding Affinity of Eukaryotic Initiation Factor 4F and (iso)4F for Cap Analogues” Wei et al. (1998) Biochem. 37:1910-16: Column 1 of page 1912 showed the K<sub>d</sub> for various cap associated proteins for PABP to be between 15nM and 40nM. (Exhibit F).

Given that the literature clearly documented the K<sub>d</sub>s of RNA binding proteins at the filing date of this application and methods of easily determining same, Applicants respectfully request that the rejection be reconsidered and withdrawn.

**Rejections Under 35 U.S.C. §102**

The Examiner rejected claims 1, 2, 8, 11-14, 17, 26, and 29 under 35 U.S.C. §102(b) as being anticipated by Allen et al. (1998) Mol. Cellul. Biol. 18:6014-6022 (“Allen”). Applicants traverse the rejection to the extent it is maintained over the claims as amended.

Applicants submit that Allen does not describe contacting a sample comprising a plurality of RNA-protein complex with at least one ligand and separating a plurality of RNP complexes by binding the ligand with a binding molecule specific for the ligand, wherein the binding molecule is attached to a solid support, collecting the RNP complexes by removing the RNP complexes from the solid support, and identifying the plurality of RNAs or other components associated with the RNP complexes. Allen describes immunoprecipitation of mRNPs using MAbs followed by either elution of proteins and protein analysis or by RNA analysis. Allen does not disclose collecting a plurality of RNP complexes by removing RNP complexes from a solid support. Allen also does not disclose identifying a plurality of RNAs or other components associated with the RNP complexes. Because Allen does not identically disclose Applicants' claimed invention, Applicants respectfully submit that Allen is not a proper reference under 35 U.S.C. §102(b). Applicants respectfully request reconsideration and withdrawal of the rejection.

The Examiner rejected claims 1, 2, 5-8, 12-15, 17, 25, and 26 under 35 U.S.C. §102(a) as being anticipated by Antic et al. (1999) Genes & Development 13:449-461 ("Antic"). Applicants traverse the rejection to the extent it is maintained over the claims as amended.

Applicants submit that Antic does not describe contacting a sample comprising a plurality of RNA-protein complexes with at least one ligand and separating the RNP complexes by binding the ligand with a binding molecule specific for the ligand, wherein the binding molecule is attached to a solid support, collecting the RNP complexes by removing the RNP complexes from the solid support, and identifying the plurality of RNAs or other components associated with the RNP complexes. Antic describes immunoprecipitation of mRNPs using an antibody to Hel-N1 protein followed by either protein analysis or RNA analysis. Antic does not disclose collecting a plurality of RNP complexes by removing an RNP complexes from a solid support. Antic also does not disclose identifying the plurality of RNAs or other components associated with the RNP complexes. Because Antic does not identically disclose Applicants' claimed invention, Applicants respectfully submit that Antic is not a proper reference under 35 U.S.C. §102(a). Applicants respectfully request reconsideration and withdrawal of the rejection.

The Examiner rejected claims 1, 2, 8, 12-14, 17, 26, and 27 under 35 U.S.C. §102(a) as being anticipated by Reim et al. (1999) Exp. Cell Res. 253:573-86 ("Reim"). Applicants traverse the rejection to the extent it is maintained over the claims as amended.

Applicants submit that Reim does not describe contacting sample comprising a plurality of RNA-protein complexes with at least one ligand and separating the RNP complexes by binding the ligand with a binding molecule specific for the ligand, wherein the binding molecule is attached to a solid support, collecting the RNP complexes by removing the RNP complexes from the solid support, and identifying the plurality of RNAs or other components associated with the RNP complexes. Reim describes immunoprecipitation of NonA protein followed by either elution of the immunocomplexes for protein analysis or elution of the immunocomplexes for RNA analysis. Reim does not disclose collecting a plurality of RNP complexes by removing an RNP complex from a solid support. Reim also does not disclose identifying the plurality of RNAs or other components associated with the RNP complexes. Because Reim does not identically disclose Applicants' claimed invention, Applicants respectfully submit that Reim is not a proper reference under 35 U.S.C. §102(b). Applicants respectfully request reconsideration and withdrawal of the rejection.

The Examiner rejected claims 1-9, 13-17, 23, and 25-28 under 35 U.S.C. §102(b) as being anticipated by Keene et al. (U.S. Patent No. 5,773,246) ("Keene"). Applicants traverse the rejection to the extent it is maintained over the claims as amended.

Keene does not describe contacting a sample comprising a plurality of RNA-protein complexes with at least one ligand and separating the RNP complexes by binding the ligand with a binding molecule specific for the ligand, wherein the binding molecule is attached to a solid support, collecting the RNP complexes by removing the RNP complexes from the solid support, and identifying the plurality of RNAs or other components associated with the RNP complexes. For example, column 24, lines 34-63 of Keene describes the binding of an RBP, Hel-N1 protein, to anti-g10 antibody and immunoprecipitation, followed by the addition of labeled RNA transcripts and isolation of RNAs that bind to the RBP. In that disclosure, an RBP was immunoprecipitated (column 24, line 41), not an mRNP complex. In the HeLa experiments

described in column 24, lines 54-63, Applicants submit that Hela nuclear extracts were used to make labeled proteins. Column 27, line 27 through column 28, line 20, which was cited by the Examiner, describes immunoprecipitation of RNPs using an anti-Hel-N1 antibody followed by extraction of RNA. Keene does not disclose collecting a plurality of RNP complexes by removing the RNP complexes from a solid support. Keene also does not disclose identifying the plurality of RNAs or other components associated with the RNP complexes. Because the Keene patent does not identically disclose Applicants' claimed invention, Applicants respectfully submit that Keene is not a proper reference under 35 U.S.C. §102(b). Applicants therefore respectfully request that the rejection be reconsidered and withdrawn.

The Examiner rejected claims 1-6, 12-15, 17, and 26 under 35 U.S.C. §102(b) as being anticipated by Buckanovich et al. (1997) Mol. and Cellul. Biol. 17(6):3194-3201 ("Buckanovich"). Applicants traverse the rejection to the extent it is maintained over the claims as amended.

Buckanovich does not describe contacting a sample comprising a plurality of RNA-protein complexes with at least one ligand and separating the RNP complexes by binding the ligand with a binding molecule specific for the ligand, wherein the binding molecule is attached to a solid support, collecting the RNP complexes by removing the RNP complexes from the solid support, and identifying the plurality of RNAs or other components associated with the RNP complexes. Buckanovich discloses immunoprecipitation of mRNPs from mouse brains followed by mRNA preparation and characterization by RT-PCR. Buckanovich does not disclose collecting a plurality of RNP complexes by removing the RNP complexes from a solid support. Buckanovich also does not disclose identifying the plurality of RNAs or other components associated with the RNP complexes. Because Buckanovich does not identically disclose Applicants' claimed invention, Applicants respectfully submit that Buckanovich is not a proper reference under 35 U.S.C. §102(b). Applicants therefore respectfully request that the rejection be reconsidered and withdrawn.

The Examiner rejected claims 1, 2, 8, 10, 18, and 26 under 35 U.S.C. §102(a) as being anticipated by Takeda et al. (1999) J. Immunol. 163:6269-6274 ("Takeda"). Applicants traverse the rejection to the extent it is maintained over the claims as amended.

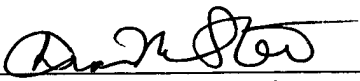
Takeda does not describe contacting a sample comprising a plurality of RNA-protein complex with at least one ligand and separating the RNP complexes by binding the ligand with a binding molecule specific for the ligand, wherein the binding molecule is attached to a solid support, collecting the RNP complexes by removing the RNP complexes from the solid support, and identifying the plurality of RNAs or other components associated with the RNP complexes. Takeda discloses immunoprecipitation and analysis of proteins or RNAs from HeLa cell extracts. Takeda does not disclose collecting a plurality of RNP complexes by removing the RNP complexes from a solid support. Takeda also does not disclose identifying the plurality of RNAs or other components associated with the RNP complexes. Because Takeda does not identically disclose Applicants' claimed invention, Applicants respectfully submit that Takeda is not a proper reference under 35 U.S.C. sec. 102(a). Applicants therefore respectfully request that the rejection be reconsidered and withdrawn.

### **CONCLUSION**

Applicants respectfully urge that all claims are in condition for allowance and request prompt and favorable action on the instant application. If the Examiner believes that a telephonic interview with the undersigned would expedite prosecution of this application, the Examiner is cordially invited to call the undersigned at (617) 338-2952.

Respectfully submitted,

Date: December 28, 2005  
Reg. No. 43,153  
Tel. No. (617) 338-2952  
Fax No. (617) 338-2880

  
\_\_\_\_\_  
Diana M. Steel, D. Phil.  
Attorney for Applicants  
Sullivan & Worcester LLP  
One Post Office Square  
Boston, MA 02109



**FMR1 Protein: Conserved RNP Family Domains and Selective RNA Binding**

Claude T. Ashley Jr.; Keith D. Wilkinson; Daniel Reines; Stephen T. Warren

*Science*, New Series, Vol. 262, No. 5133 (Oct. 22, 1993), 563-566.

Stable URL:

<http://links.jstor.org/sici?sici=0036-8075%2819931022%293%3A262%3A5133%3C563%3AFPCRFD%3E2.0.CO%3B2-S>

*Science* is currently published by American Association for the Advancement of Science.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/aaas.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

tion of antagonistic muscles has been demonstrated as the locust escape jump [R. H. J. Brown, *Nature* 214, 939 (1967)] and the predatory strike of the dragonfly larva [Y. Tanaka and M. Hisada, *J. Exp. Biol.* 88, 1 (1980)] or of the mantid shrimp [M. Burrows, see (7)].

19. We thank S. O. Andersen for his advice suggesting that the tendon between the mandible and its closer muscle would be the most probable region in which to find energy-storing materials such as resilin. We also thank the German Institute for the Scientific Film (IWF), where the high-speed cinematography was carried out. Supported by funds from the Deutsche Forschungsgemeinschaft (Leibniz Preis to B.H., SFB 251/project 18, and Gr 933/3).

15 June 1993; accepted 12 August 1993

## FMR1 Protein: Conserved RNP Family Domains and Selective RNA Binding

Claude T. Ashley Jr., Keith D. Wilkinson, Daniel Reines, Stephen T. Warren\*

Fragile X syndrome is the result of transcriptional suppression of the gene *FMR1* as a result of a trinucleotide repeat expansion mutation. The normal function of the *FMR1* protein (FMRP) and the mechanism by which its absence leads to mental retardation are unknown. Ribonucleoprotein particle (RNP) domains were identified within FMRP, and RNA was shown to bind in stoichiometric ratios, which suggests that there are two RNA binding sites per FMRP molecule. FMRP was able to bind to its own message with high affinity (dissociation constant = 5.7 nM) and interacted with approximately 4 percent of human fetal brain messages. The absence of the normal interaction of FMRP with a subset of RNA molecules might result in the pleiotropic phenotype associated with fragile X syndrome.

Fragile X syndrome is an X-linked dominant disorder with reduced penetrance that occurs at a frequency of approximately 0.5 to 1.0 per 1000 males and 0.2 to 0.6 per 1000 females (1). Fully penetrant males exhibit moderate mental retardation along with a phenotype consisting of macroorchidism (enlarged testes), subtle facial dysmorphism, and mild connective tissue abnormalities (2). Female patients typically are less severely affected, showing little or no somatic signs and only borderline to mild mental retardation. The molecular basis of fragile X syndrome has been attributed to the expansion of an unstable CGG trinucleotide repeat in the 5' untranslated region of the gene *FMR1* (3, 4).

In fragile X syndrome, when the size of the CGG repeat is in the affected range beyond 230 repeats, the *FMR1* gene is methylated; this methylation results in transcriptional silencing (5). The absence of *FMR1* message and its encoded protein, FMRP, is believed responsible for the phenotype of the fragile X syndrome. In addition to the common mutational change of repeat expansion, three variant patients with the clinical presentation of fragile X syndrome have been reported: two males with large deletions encompassing the *FMR1* locus and a severely affected male with a *FMR1* Ile<sup>304</sup>→Asn missense mutation (6).

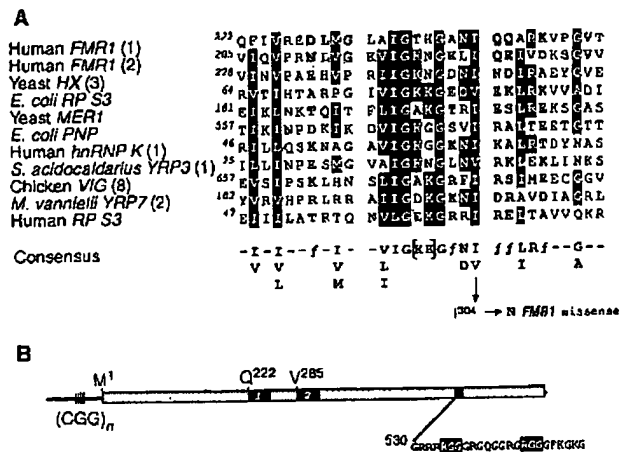
Alternative splicing generates several isoforms of FMRP with a major species of 69 kD (7). Although predominantly cytoplasmic, occasional nuclear localization is observed (8). In situ hybridization with *FMR1* mRNA reveals widespread but not ubiquitous expression with abundant message present in the testes and in neurons in the brain (9).

Initial analyses of *FMR1* and the pre-

dicted protein sequence revealed little sequence similarity to known proteins or motifs. Further analyses of the human and mouse genes, particularly with the use of searches of limited regions of 200 to 400 amino acids in length, revealed two similar regions of FMRP that also were similar to 6 repetitive domains in the yeast protein HX and 14 domains of the chicken gene vigilin (VIG) (10). Alignments of these amino acid sequences and a resulting profile search revealed a number of proteins containing 1 to 14 repeats of an uninterrupted, 30-amino acid domain (Fig. 1A). Proteins containing this domain, termed KH domains, are believed to constitute a ribonucleoprotein (RNP) family (11) that includes mer1, a yeast protein involved in meiosis-specific alternative splicing (12); bacterial polynucleotide phosphorylase, which binds RNA and has phosphorolysis activity (13); and the highly conserved ribosomal protein S3 (RP S3) (14). Thus, most functional aspects of RNA-protein interactions are represented among KH domain-containing proteins, including RNA catalysis, message processing, and translation.

The two KH domains of FMRP reside in the middle of the protein (Fig. 1B), a

**Fig. 1.** Location and homologies of RNP family domains in FMRP. (A) Alignment (27) of the amino acid sequences that make up the KH domains of FMRP and several other proteins and the corresponding consensus sequence. Numbers in parentheses indicate the particular domain shown for the proteins that have multiple KH domains, and the number preceding the first residue indicates that position in the corresponding protein. Dark highlighting indicates similarities among all proteins, whereas stippled highlighting indicates similarity between the two KH domains of FMRP. Boldface residues show the positions of polar amino acids, indicated by / in the consensus sequence. The bracketed lysine (K) residues indicate this amino acid at either position in the domain. The position of the isoleucine-to-asparagine mutation at position 304 (I<sup>304</sup>→N) in a patient (6) is indicated at the bottom. (B) Diagram of FMRP [residue numbers are as described (7)]. The CGG repeat and initiating codon (M<sup>1</sup>) are indicated as is each KH domain, labeled 1 and 2. Also shown is the amino acid sequence with the two RGG box domains highlighted. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.



C. T. Ashley Jr., K. D. Wilkinson, D. Reines, Department of Biochemistry, Emory University School of Medicine, Atlanta, GA 30322.

S. T. Warren, Howard Hughes Medical Institute and Departments of Biochemistry and Pediatrics, Emory University School of Medicine, Atlanta, GA 30322.

\*To whom correspondence should be addressed.

region not involved in alternative splicing (7). Furthermore, the two FMRP domains show 100% amino acid identity among the human, mouse, and chicken FMR1 homologs (15). In the single patient with severe fragile X syndrome that was a result of a missense FMR1 mutation (6), this mutation was in an invariant isoleucine in the 20th position of the second KH domain (Fig. 1A). In addition, FMRP shows two RGG (Arg-Gly-Gly) boxes toward the carboxyl end. RGG boxes are another distinct RNP motif found in heterogeneous nuclear RNP K (hnRNP K) and nucleolar proteins (16).

To biochemically establish and characterize RNA binding by FMRP, we measured <sup>35</sup>S-labeled FMRP bound to biotinylated RNA (17). The <sup>35</sup>S-FMRP was translated in vitro from a full-length FMR1 complementary DNA (cDNA) (18), which resulted in a 69-kD protein as well as smaller, secondary products of the reaction (Fig. 2A, lane 1). FMR1 RNA was transcribed in vitro from the same cDNA with the T3 (sense) or T7 (antisense) promoter in the presence of biotin-uridine 5'-triphosphate (UTP) (19). Biotinylated RNA was mixed with the in vitro translation reaction and captured with streptavidin-linked magnetic beads. After three washes, bound protein was released by denaturation and assayed by SDS-polyacrylamide gel electrophoresis (PAGE) and fluorography. A known RNA binding protein, Brome mosaic virus (BMV) RNA-dependent RNA polymerase, was captured by this assay (Fig. 2A), whereas no binding was observed with a fourfold higher molar concentration of the 15-kD parathyroid hormone precursor protein, which does not associate with RNA (20). Only the 69-kD protein, corresponding to the mass of FMRP, was captured from the FMR1 in vitro translation mix by the biotinylated RNA. Hydroxylamine cleavage of this 69-kD protein, released from the biotinylated RNA, resulted in a doublet band

at approximately 34 kD. Two bands of 33 and 34 kD are predicted based on the FMRP sequence, which confirms that this species is FMRP (21).

The FMRP was able to bind to sense or antisense FMR1 RNA (Fig. 2B). Experiments were performed in the presence of a 62-fold excess of transfer RNA, suggesting that the protein specifically binds to unstructured RNA, perhaps allowing access to the sugar-phosphate backbone and unpaired bases for protein interaction. Ribonuclease (RNase) treatment before streptavidin exposure as well as use of nonbiotinylated RNA resulted in no capture of FMRP (Fig. 2B). FMRP binds strongly to single-stranded DNA but to a lesser extent to double-stranded DNA (Fig. 2B), which is similar to the binding of other RNP family proteins (22). To test whether FMRP binds to all RNA species, we pooled a human fetal brain plasmid cDNA library (23) and performed in vitro transcription in the presence of biotin-UTP. A complex mixture of RNAs (average size of 700 base pairs) was produced, some of which allowed the capture of FMRP (Fig. 2B, lane 4). However, much less RNA binding was observed with the pooled RNAs as compared with that in the presence of FMR1 RNA, even with a sixfold molar excess of the former over the latter. Thus, it appears that FMRP is not a general, nonspecific RNA binding protein but rather exhibits some degree of specificity for the RNA species with which it interacts.

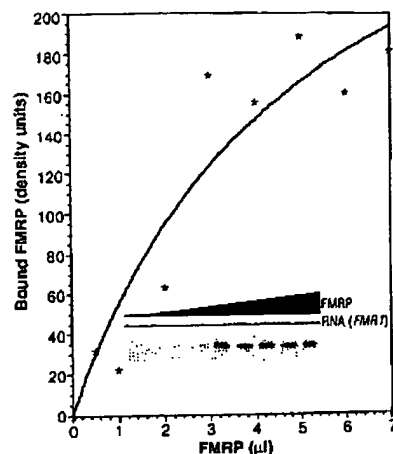
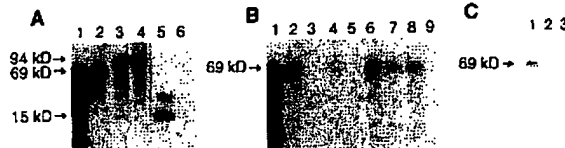
As an independent measure of FMRP binding to RNA, a monoclonal antibody (D44), which binds specifically to RNA (24), was used (Fig. 2C). D44 immunoprecipitation of nonbiotinylated RNA with bound FMRP was achieved when protein A-linked agarose beads were used to capture the complex. RNase treatment abolished FMRP capture by this assay as well.

Estimates of the strength of the interaction of FMRP with RNA were next ob-

tained. Protein titration studies (Fig. 3) demonstrate saturation behavior of FMRP binding to 80 ng of RNA. At saturation, 3 nM FMRP was bound, as estimated by direct scintillation counting of the bound FMRP (25). Because the RNA concentration was held constant at 6 nM, a 2:1 stoichiometry of RNA:protein was obtained. Similar studies with RNA from human fetal brain cDNA transcribed in vitro also displayed saturation, but only 1.5 nM FMRP was bound. Considering that the brain RNA was held at a higher concentration of 34 nM, a stoichiometry of approximately 27:1 (RNA:protein) was determined, which suggests that there was selective FMRP binding to approximately 4% by mass of the human fetal brain message. Because FMR1 is estimated to account for less than 0.01% of this pool, these data demonstrate FMRP interaction with messages of other human genes.

Biotinylated RNA transcripts from randomly chosen human fetal brain cDNA clones were tested for selective FMRP binding. Analysis of 12 clones identified two that produced RNA that bound to FMRP (Fig. 4). Insert size appeared unrelated to binding ability, and preliminary sequence analysis of those clones whose RNA bound did not reveal any obvious homology to FMR1 cDNA nor to any database sequence. Thus, selective RNA binding of

**Fig. 2.** Binding of RNA to FMRP. (A) Validation of the biotinylated RNA binding assay (28) with in vitro-translated FMRP (lanes 1 and 2). BMV RNA-dependent polymerase (positive control, lanes 3 and 4), and parathyroid hormone precursor protein (negative control, lanes 5 and 6). The complete in vitro translation mix of each is shown before (lanes 1, 3, and 5) and after (lanes 2, 4, and 6) binding to in vitro-transcribed biotinylated FMR1 RNA (19). (B) Nucleic acid binding properties of 2  $\mu$ l of in vitro-translated FMRP (lane 1) with 80 ng of various nucleic acids: lanes 2 and 3, biotinylated and nonbiotinylated, respectively, FMR1 RNA (sense); lanes 4 and 5, biotinylated and nonbiotinylated, respectively, human fetal brain RNA pool (23); lane 6, biotinylated FMR1 RNA (antisense); lanes 7 and 8, double- and single-stranded, respectively, FMR1 DNA (29); and lane 9, biotinylated FMR1 RNA (sense) treated with RNase (1  $\mu$ g/ $\mu$ l of RNase A and 1 U of T1 RNase) before streptavidin capture. (C) Co-immunoprecipitation of 2  $\mu$ l of in vitro-translated FMRP and nonbiotinylated FMR1 RNA (30) in the presence (lane 1) or absence (lane 2) of the D44 RNA antibody (25). Lane 3, precipitated FMRP in the presence of antibody and RNase (1  $\mu$ g/ $\mu$ l of RNase A and 1 U of RNase T1).



**Fig. 3.** Titration of FMRP with a constant amount of biotinylated FMR1 RNA. Amounts of in vitro-translated FMRP ranging from 0 to 7  $\mu$ l were incubated with 80 ng of biotinylated FMR1 RNA, and captured protein was eluted and analyzed as described. The fluorogram of the individual binding reactions (inset) was subjected to densitometry, and amounts of bound protein inferred from the density of the individual bands (y axis) were plotted versus the amount of in vitro-translated FMRP (x axis) used (26). Asterisks denote individual data points.

FMRP is substantiated, and further analysis of these clones, in conjunction with analysis of *FMRI*, should allow recognition of any structural RNA determinants required for FMRP binding.

As would be predicted for saturable binding of FMRP to RNA, competition was observed with nonbiotinylated RNA (Fig. 5A). Saturable binding was further substantiated by a series of analyses in which the amount of labeled FMRP was kept constant

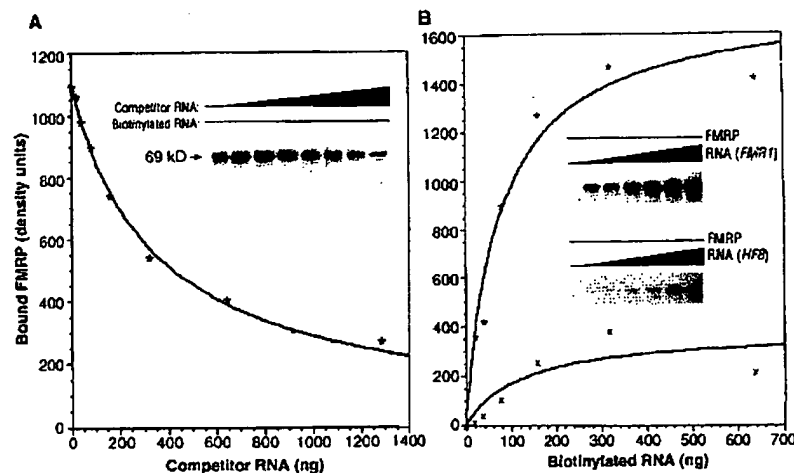
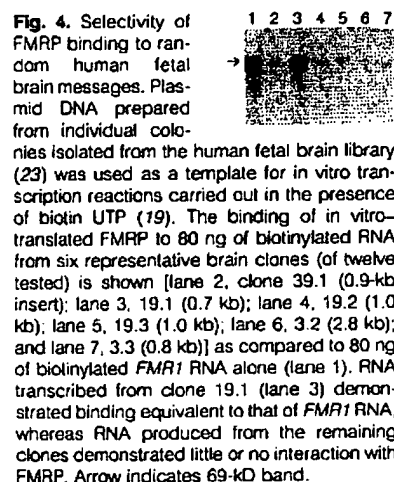
and the amount of biotinylated *FMRI* RNA or biotinylated brain RNA was increased (Fig. 5B). Curve fitting with a simple binding equation (26) resulted in an apparent dissociation constant ( $K_d$ ) of 5.7 nM for *FMRI* RNA. When human fetal brain RNA was used, an apparent  $K_d$  of 39 nM was obtained. These results were confirmed by direct scintillation counting of bound  $^{35}$ S-FMRP. Substitution of the value for the estimated  $K_d$  (5.7 nM) into the equation that described the competition data (Fig. 5A) produced an estimated inhibition constant ( $K_i$ ) of 12.4 nM. The difference in estimated values for  $K_d$  and  $K_i$  may reflect either that the biotinylated RNA acts as a better ligand than nonbiotinylated RNA or that there are two RNA binding sites per protein molecule. The latter possibility would reduce the apparent binding constant because biotinylated RNA bound to either or both sites would result in capture of the FMRP molecule by streptavidin. Because the stoichiometry at saturation suggested a 2:1 *FMRI* RNA:FMRP ratio, this possibility is likely. One may also speculate that either the KH or RGG domains, each present twice in FMRP, are directly binding RNA.

The demonstration that FMRP is an RNA binding protein suggests an avenue of further inquiry as to the precise function of FMRP. Given that the major neurological

phenotype of fragile X syndrome is mental retardation, these data provide a means to explore the biochemical basis of cognitive function in humans as well as the pathophysiology of fragile X syndrome. The observation that FMRP interacts with a selective, but substantial, fraction of human brain RNA suggests that other genes, whose products may vary in the absence of FMRP, could play a consequential role in the pleiotropic phenotype of fragile X syndrome.

## REFERENCES AND NOTES

- W. T. Brown, *Am. J. Hum. Genet.* 47, 175 (1990); S. L. Sherman *et al.*, *Ann. Hum. Genet.* 48, 21 (1984).
- M. G. Butler, T. Mangrum, R. Gupta, D. N. Singh, *Clin. Genet.* 39, 347 (1991).
- A. M. J. H. Verkerk *et al.*, *Cell* 65, 905 (1991); I. Oberlé *et al.*, *Science* 252, 1097 (1991); E. J. Kremer *et al.*, *ibid.*, p. 1711; A. Vincent *et al.*, *Nature* 349, 624 (1991).
- Y. H. Fu *et al.*, *Cell* 67, 1047 (1991); S. Yu *et al.*, *Science* 252, 1179 (1991).
- M. Pieretti *et al.*, *Cell* 66, 817 (1991); J. S. Sulcliffe *et al.*, *Hum. Mol. Genet.* 1, 397 (1992).
- D. Wahnle *et al.*, *Am. J. Hum. Genet.* 51, 299 (1992); A. K. Gideon *et al.*, *Nature Genet.* 1, 341 (1992); K. De Boulle *et al.*, *ibid.* 3, 31 (1993).
- C. T. Ashley *et al.*, *Nature Genet.* 4, 244 (1993).
- C. Verheij *et al.*, *Nature* 363, 722 (1993).
- H. L. Hinds *et al.*, *Nature Genet.* 3, 36 (1993).
- A. Delahodde, A. M. Becam, J. Perea, C. Jacq, *Nucleic Acids Res.* 14, 9213 (1986); C. Schmidt *et al.*, *Eur. J. Biochem.* 208, 625 (1992).
- H. Stoml, M. J. Matusis, W. M. Michael, G. Dreyfuss, *Nucleic Acids Res.* 21, 1193 (1993).
- J. Engebrecht and G. S. Roeder, *Mol. Cell. Biol.* 10, 2379 (1990).
- P. Regnier, M. Grunberg-Manago, C. Portier, *J. Biol. Chem.* 262, 63 (1987).
- X. T. Zhang, Y.-M. Tan, Y. H. Tan, *Nucleic Acids Res.* 18, 6689 (1990); D. Brauer and R. Roring, *FEBS Lett.* 106, 352 (1979).
- D. Price and S. T. Warren, unpublished material.
- I. W. Mattaj, *Cell* 73, 837 (1993).
- W. C. Boelens *et al.*, *ibid.* 72, 881 (1993).
- The full-length murine *FMRI* cDNA (4256 base pairs) Mc 2.17 described in Ashley *et al.* (7) was linearized for transcription with either *Nsi*I (sense) or *Not*I (antisense). In vitro transcription was carried out with the T3-T7 in vitro transcription kit (Stratagene). In vitro translation was in rabbit reticulocyte lysates of the In Vitro Express kit (Stratagene).
- Biotinylated RNA (*FMRI* and *HFB*) was synthesized in the presence of biotinylated ribonucleoside triphosphate (rUTP) (ENZO, Farmingdale, NY) as described (17) at a molar ratio of 1:10 (biotinylated rUTP:rUTP).
- BMV RNA-dependent RNA polymerase [C. C. Kao, R. Quadt, R. P. Hersherberger, P. Ahlquist, *J. Virol.* 66, 6322 (1992)] was synthesized from RNA purchased from Promega. Only the 94-kD RNA-dependent RNA polymerase was translated in the In Vitro Express system (Stratagene). Parathyroid hormone precursor was translated from RNA included in the In Vitro Express system.
- Captured protein from four standard binding reactions (80 ng of biotinylated *FMRI* RNA and 2  $\mu$ l of in vitro-translated FMRP) was resuspended in 100  $\mu$ l of 2 M hydroxylamine HCl in 6 M guanidine HCl (pH 9.0). After a 90-min incubation at 45°C, the cleaved protein was released by heating (3 min at 100°C) and analyzed by SDS-PAGE, as described (28). Predicted hydroxylamine cleavage sites at residues 13, 301, and 612 of the FMRP sequence (7) resulted in a doublet band at 34 kD.
- M. J. Matusis, W. M. Michael, G. Dreyfuss, *Mol. Cell. Biol.* 12, 164 (1992).



**Fig. 5.** Characterization of FMRP binding to *FMRI* RNA. (A) Competition of FMRP binding by addition of nonbiotinylated *FMRI* RNA. Constant amounts of FMRP (2  $\mu$ l) and biotinylated *FMRI* RNA (80 ng) were incubated with increasing amounts of nonbiotinylated *FMRI* RNA (0 to 1280 ng), and captured FMRP was analyzed by SDS-PAGE and fluorography (inset). Densitometry was used to estimate amounts of bound protein, and bound FMRP (y axis) was plotted (26) against amounts of nonbiotinylated *FMRI* RNA (x axis). Asterisks denote individual data points. (B) RNA titrations of FMRP binding with increasing amounts of *FMRI* RNA or human fetal brain (HFB) RNA. A constant amount of FMRP (2  $\mu$ l) was incubated in the presence of increasing amounts of either biotinylated *FMRI* RNA (asterisks) or biotinylated HFB RNA (x) ranging from 0 to 640 ng, and the FMRP captured was analyzed by SDS-PAGE followed by fluorography. Fluorograms (insets) were scanned by densitometry, and the relative amount of bound FMRP (y axis) was plotted (26) versus the amount of biotinylated RNA added (x axis).

23. The human fetal brain plasmid cDNA library (Invitrogen) was amplified in toto, and cesium-purified plasmid DNA was linearized at the Xba I site. In vitro transcription was performed from a T7 promoter as described (18, 19).
24. D. Eilat, M. Hochberg, R. Fischel, R. Laskov, *Proc. Natl. Acad. Sci. U.S.A.* 79, 3818 (1982).
25. The specific activity of [<sup>35</sup>S]methionine used was 1275 Ci/mmol. Given the molecular weight (69 kD) and that nine methionines existed in FMRP (7), the expected number of disintegrations per minute per mole of protein was calculated and adjusted according to the predetermined efficiency of counting (88%) for the isotope. The number of moles of bound protein was determined by dividing the observed counts per minute of bound protein at saturation by the expected disintegrations per minute per mole; the concentration value given was determined by dividing by the reaction volume.
26. The standard binding equation used was  $b = \frac{b_m([L]/K_d + [L])}{1 + [L]/K_d}$ , where  $b$  is the amount of protein bound,  $b_m$  is the maximum amount bound,  $[L]$  is the concentration of ligand, and  $K_d$  is the dissociation constant. Maximum amounts of bound protein were determined from direct scintillation counting of binding reactions performed in triplicate at constant amounts of ligand. Amounts of bound protein were determined either through direct scintillation counting or densitometric analysis of fluorograms. Data points were fit to this equation with the nonlinear least squares method furnished in the plotting program Delta Graph (Deltapoint, Inc., Monterey, CA), and equations were solved for the apparent  $K_d$ . A variation of this equation,  $b = \frac{b_m([L]/K_d + [L] + [I]/K_i)}{1 + [L]/K_d + [I]/K_i}$  (where  $[I]$  = concentration of inhibitor), was used to fit the RNA competition curve and to solve for  $K_i$ .
27. M. Gribskov, R. Luethy, D. Eisenberg, *Methods Enzymol.* 183, 146 (1989).
28. Biotinylated RNA binding assays were performed as described by Boelens *et al.* (17), except that magnetic beads with conjugated streptavidin (Dynal) were used. Captured material was resuspended in 20  $\mu$ l of 1 $\times$  SDS sample buffer, and bound protein was eluted by boiling for 10 min and resolved by SDS-PAGE. The 12% SDS-polyacrylamide gels were soaked for 30 min in destain solution (7.5% methanol and 10% acetic acid), followed by 1 hour in 1 M salicylate solution (1 M salicylate, 30% methanol, and 3.0% glycerol). Gels were then dried and exposed at -80°C with an intensifying screen for 24 to 72 hours.
29. Biotinylated DNA was prepared with a nick translation kit (Amersham) in the presence of biotinylated deoxythymidine triphosphate (dTTP) (BRL) at an equal molar ratio with dTTP and passed through a G-50 Sephadex spin column (Boehringer Mannheim) followed by precipitation. Single-strand DNA was obtained by heat denaturation of the product above.
30. Immunoprecipitations were carried out in the presence of 2  $\mu$ l of in vitro-translated FMRP, D44 (20 ng/ $\mu$ l), and 80 ng of nonbiotinylated FMRP RNA by means of the same protocol as used in the biotinylated RNA binding assays (17). RNA molecules bound to antibody and FMRP were captured with protein A that had been linked to agarose beads (BRL; amount equivalent to 20  $\mu$ l). After centrifugation, pellets were washed as described (17), resuspended in 20  $\mu$ l of 1 $\times$  SDS buffer, and resolved by SDS-PAGE.
31. Supported by NIH grant HD20521 (S.T.W.). C.T.A. is a predoctoral fellow of the March of Dimes Birth Defects Foundation, and S.T.W. is an Investigator of the Howard Hughes Medical Institute.

21 July 1993; accepted 30 August 1993

## A Yeast Protein Similar to Bacterial Two-Component Regulators

Irene M. Ota and Alexander Varshavsky\*

Many bacterial signaling pathways involve a two-component design. In these pathways, a sensor kinase, when activated by a signal, phosphorylates its own histidine, which then serves as a phosphoryl donor to an aspartate in a response regulator protein. The Slr1 protein of the yeast *Saccharomyces cerevisiae* has sequence similarities to both the histidine kinase and the response regulator proteins of bacteria. A missense mutation in *SLN1* is lethal in the absence but not in the presence of the N-end rule pathway, a ubiquitin-dependent proteolytic system. The finding of *SLN1* demonstrates that a mode of signal transduction similar to the bacterial two-component design operates in eukaryotes as well.

In bacteria, a broad spectrum of responses to an often rapidly changing environment is mediated by mechanistically similar pathways known as two-component systems. The functions of two-component pathways include chemotaxis, sporulation, osmoregulation, transformation competence, virulence, and responses to changes in the sources of carbon, nitrogen, oxygen, and phosphorus (1-3). The sensor component of these systems is often an integral membrane protein containing a cytosolic trans-

mitter domain that acts as a histidine-phosphorylating autokinase, when activated by a specific signal. The signal is sensed by a distinct input domain, often located in the periplasmic space. The histidine-linked phosphoryl group of an activated transmitter is transferred to an aspartate in the receiver domain of a response regulator protein, the second component of the pathway. Receiver phosphorylation regulates the activity of its output module, which is often a DNA-binding domain (1-3).

Although two-component pathways are common in bacteria, evidence for their existence in eukaryotes has been scarce. A protein kinase from rat mitochondria has

sequence similarities to bacterial histidine kinases; however, in vitro it phosphorylates a Ser residue (4). Another eukaryotic candidate is phytochrome, a plant regulatory protein that has weak but potentially significant similarities to the sequences of bacterial histidine kinases (5). Histidine kinase activity has been detected in extracts from *S. cerevisiae* and other eukaryotic cells (6).

This report describes the *S. cerevisiae* *SLN1* gene which encodes a 134-kD product with strong sequence similarities to both the transmitter and receiver domains of the bacterial two-component regulators. We found *SLN1* while studying the N-end rule, a relation between the in vivo half-life of a protein and the identity of its N-terminal residue (7, 8). The N-end rule is a consequence of a set of ubiquitin-dependent degradation signal called N-degrons (9). Ubiquitin is a protein whose covalent conjugation to other proteins plays a role in a number of cellular processes, primarily through routes that involve protein degradation (8). In *S. cerevisiae*, the recognition component of the N-end rule pathway is encoded by the *UBR1* gene (10). A *ubr1 $\Delta$*  mutant is viable, grows at nearly wild-type rates, but is unable to degrade N-end rule substrates, which are short-lived in wild-type (*UBR1*) cells (8, 10).

In a search for the functions of the N-end rule, we carried out a "synthetic lethal" screen to identify mutants whose viability requires the presence of the *UBR1* gene (11). To isolate such mutants, termed *sin* (synthetic lethal of N-end rule), we used a screen based on 5-fluoroorotic acid (FOA) (12). In this method, yeast cells lacking chromosomal copies of both *URA3* and a (nonessential) gene of interest are transformed with a plasmid that expresses both of these genes. The cells are mutagenized and examined for growth on plates containing FOA and uracil. Because FOA selects against *URA3*-expressing cells (12), mutants that grow in the absence but not in the presence of FOA should include those whose viability requires the plasmid carrying the gene of interest linked to *URA3*.

A synthetic lethal screen with *UBR1* yielded a recessive mutant, *sln1-1*, which was viable in the *UBR1* but not in the *ubr1 $\Delta$*  background (11). Our earlier attempts to clone *SLN1* yielded PTP2, which encodes a putative phosphotyrosine phosphatase, and is an extragenic multicopy suppressor of *sln1-1* (11). The *SLN1* gene was isolated as described (13). That the subcloned DNA fragment containing a single open reading frame (ORF) was indeed *SLN1* (rather than an extragenic suppressor of *sln1-1*) was confirmed by linkage analysis (13). The 3.66-kb ORF of *SLN1* encodes a

Division of Biology, California Institute of Technology, Pasadena, CA 91125.

\*To whom correspondence should be addressed.

# AUF1 Binding Affinity to A+U-rich Elements Correlates with Rapid mRNA Degradation\*

(Received for publication, January 31, 1996, and in revised form, February 29, 1996)

Christine T. DeMaria† and Gary Brewer§

From the Department of Microbiology and Immunology, Bowman Gray School of Medicine of Wake Forest University, Winston-Salem, North Carolina 27157-1064

Rapid degradation of many labile mRNAs is regulated in part by an A+U-rich element (ARE) in their 3'-untranslated regions. Extensive mutational analyses of various AREs have identified important components of the ARE, such as the nonamer motif UUAUUUAUU, two copies of which serve as a potent mRNA destabilizer. To investigate the roles of *trans*-acting factors in ARE-directed mRNA degradation, we previously purified and molecularly cloned the RNA-binding protein AUF1 and demonstrated that both cellular and recombinant AUF1 bind specifically to AREs as shown by UV cross-linking assays *in vitro*. In the present work, we have examined the *in vitro* RNA-binding properties of AUF1 using gel mobility shift assays with purified recombinant His<sub>6</sub>-AUF1 fusion protein. We find that ARE binding affinities of AUF1 correlate with the potency of an ARE to direct degradation of a heterologous mRNA. These results support a role for AUF1 in ARE-directed mRNA decay that is based upon its affinity for different AREs.

Control of mRNA stability is an important component of eukaryotic gene expression and involves *cis*-acting elements that can be found in the coding region and/or UTRs<sup>1</sup> of mRNAs (reviewed in Refs. 1–6). One type of *cis*-acting instability element is comprised of the AREs found in the 3'-UTRs of many unstable mRNAs (reviewed in Ref. 7). Many ARE-containing mRNAs are degraded by a sequential pathway involving removal of the poly(A) tract followed by degradation of the mRNA body (8–10). In most cases the poly(A) tract is thought to protect the mRNA from ribonuclease attack so that its removal permits degradation of the mRNA body (reviewed in Ref. 11). While it has been known for almost a decade that AREs are important for mRNA instability (12–15), the mechanism(s) by which they mediate mRNA turnover is still unknown.

Despite the presence of AREs in many different mRNAs, there is no single evolutionarily conserved A+U-rich instability sequence. Typically AREs contain multiple copies of the pen-

tanucleotide AUUUA, often in conjunction with one or more U-rich regions (14). In addition, transfection studies indicate that as the number of tandemly repeated AUUU motifs is increased in a reporter mRNA, its instability increases. Likewise, two copies of the nonameric motif UUAUUUAUU act as a more potent destabilizer than a single nonameric motif (16, 17). Together, these analyses suggest that potent destabilizing AREs are high affinity binding sites for a mRNA decay factor(s).

In order to investigate how AREs function in mRNA turnover, we utilized a cell-free mRNA decay system to identify proteins that may be relevant to ARE-directed mRNA decay (8, 18, 19). To this end, we previously reported the purification, molecular cloning, and characterization of the ARE-binding protein AUF1 (20). Cellular AUF1 purified from cytoplasmic extracts of K562 human erythroid leukemia cells consists of a 37- and a 40-kDa isoform. Cloning of the 37-kDa isoform, p37<sup>AUF1</sup>, revealed two nonidentical RNA recognition motifs (21) and a short glutamine-rich region in the predicted amino acid sequence. Cloning of murine cDNAs suggests that the 40-kDa isoform may also contain 19 additional amino acids N-terminal to RNA recognition motif 1 (22). Both cellular and recombinant p37<sup>AUF1</sup> (hereafter referred to as AUF1) bind the AREs present in the *c-fos* and *c-myc* proto-oncogene mRNAs and the granulocyte-macrophage colony-stimulating factor cytokine mRNA as shown by UV cross-linking assays *in vitro*.

The potential influence of AUF1 on ARE-directed mRNA decay extends beyond the control of cytokine and proto-oncogene expression, however. Many mRNAs encoding components of G protein-coupled receptors, such as  $\beta$ -adrenergic receptors ( $\beta$ -ARs), contain AREs. Moreover, receptor levels are frequently subject to regulatory control. For example, exposure of smooth muscle cells to agonist down-regulates  $\beta_2$ -AR mRNA levels by inducing degradation of the mRNA (23). Similarly, agonist-mediated destabilization of the human  $\beta_1$ -AR mRNA appears to be dependent upon an ARE (24), and for both the human  $\beta_1$ -AR and hamster  $\beta_2$ -AR mRNAs decay occurs concomitantly with an increase in the cytoplasmic levels of AUF1 (25). Since both cellular and bacterially expressed AUF1 bind the  $\beta_2$ -AR ARE (25), the reciprocal relationship between the half-life of  $\beta_2$ -AR mRNA and the abundance of AUF1 suggests that the half-lives of ARE-containing mRNAs may be dependent in part upon ARE-specific RNA binding affinity of AUF1.

Here, we test the hypothesis that the binding affinity of AUF1 for an ARE should reflect the potency of that ARE as a mRNA destabilizer. Using purified recombinant His<sub>6</sub>-AUF1 fusion protein, we find a direct relationship between the apparent  $K_d$  for ARE binding by AUF1 and the potency of the ARE to direct mRNA decay. These results support a role for AUF1 in ARE-directed mRNA decay that is based upon its affinity for different AREs.

\* This work was supported by American Cancer Society Grant NP-884 (to G. B.). Oligodeoxynucleotide synthesis was performed by the DNA Synthesis Core Laboratory of the Comprehensive Cancer Center of Wake Forest University supported in part by National Institutes of Health (NIH) Grant CA12197. PhosphorImager facilities were also supported by NIH Grant CA12197 and by North Carolina Biotechnology Center Grant 9510-IDG-1006. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Supported by NIH Training Grant T32-AI07401.

§ To whom correspondence should be addressed: Dept. of Microbiology and Immunology, Bowman Gray School of Medicine of Wake Forest University, Medical Center Blvd., Winston-Salem, NC 27157-1064. Tel.: 910-716-6756; Fax: 910-716-9928; E-mail: gbrewer@bgsu.edu.

<sup>1</sup> The abbreviations used are: UTR, untranslated region; ARE, A+U-rich element;  $\beta$ -AR,  $\beta$ -adrenergic receptor.

## MATERIALS AND METHODS

All enzymes and plasmid vectors were obtained from Promega Corp. (Madison, WI) unless otherwise noted. All plasmid constructions were confirmed by both restriction enzyme analyses and dideoxy sequencing with Sequenase (version 2.0, U.S. Biochemical Corp.).

**Expression and Purification of His<sub>6</sub>-AUF1 Fusion Protein**—His<sub>6</sub>-AUF1 fusion protein was expressed in bacteria and purified as described by Pende *et al.* (25). The concentration of purified recombinant His<sub>6</sub>-AUF1 was estimated by comparison with dilutions of BSA using Coomassie-stained SDS-polyacrylamide gels.

**Gel Filtration Analysis of His<sub>6</sub>-AUF1**—A 1 × 19-cm column of Sephacryl S-300 (Pharmacia Biotech Inc.) was equilibrated in factor-binding buffer (10 mM Tris, pH 7.5, 5.5 mM magnesium acetate, 100 mM potassium acetate). Four micromolar purified recombinant His<sub>6</sub>-AUF1 in a final volume of 300 μl of factor-binding buffer was loaded onto the column; 60 300-μl fractions were collected. Twenty microliters of each fraction was assayed for His<sub>6</sub>-AUF1 by Western blot analysis using anti-AUF1 antisera, and 5 μl of each fraction was analyzed for RNA-binding activity by mobility shift assay with radiolabeled *c-fos* ARE. Similarly, 400 nM purified recombinant His<sub>6</sub>-AUF1 in 300 μl of factor-binding buffer was loaded onto the column, and 20 900-μl fractions were collected. Protein in these fractions was precipitated with 10% trichloroacetic acid, 20 μg/ml lysozyme by incubation on ice for 30 min and centrifugation at 4 °C for 30 min. Each precipitate was resuspended in 10 mM Tris (pH 7.5) and assayed for His<sub>6</sub>-AUF1 by Western blot. The void volume was determined using blue dextran.

**Construction of Plasmids for *in Vitro* RNA Synthesis**—The sequences Rβ+AT×1, Rβ+AT×2, Rβ+AT×3, and Rβ+AT×5 were synthesized by polymerase chain reaction using plasmids pNEORβG<sup>AT1×</sup>, pNEORβG<sup>AT2×</sup>, pNEORβG<sup>AT3×</sup>, and pNEORβG<sup>AT5×</sup> (Ref. 26; gifts of Gray Shaw, Genetics Institute), respectively, as DNA templates. For each reaction, the 5' oligo primer was 5'-CTGTCTCATCTTTTGG-3', and the 3' oligo primer was 5'-CGCGGTACCGAAGAGGGACAGC-TATG-3'. Amplified fragments were digested with *KpnI* and *EcoRI* and ligated to *KpnI*-*EcoRI* digested pT7T3α19 (Life Technologies, Inc.) to create plasmids pα19Rβ+AT×1, pα19Rβ+AT×2, pα19Rβ+AT×3, and pα19Rβ+AT×5.

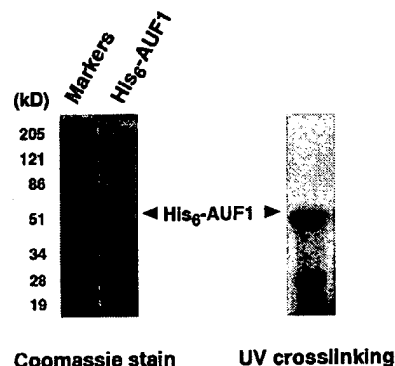
Plasmid pα19Rβ+AT×4 was synthesized by first annealing the complementary oligonucleotides 5'-CTAGATTTATTTATTTAGCTT-TAG-3' and 5'-TCGACTAAAGCTAAATAAATAAATAA-3' and then ligating to the 3.1-kilobase vector fragment derived from *XbaI*-*SaII* digested pα19Rβ+AT×5.

The U<sub>32</sub> sequence was created as a *Bam*HI-*Bgl*II fragment by annealing the complementary oligonucleotides 5'-GATCCT<sub>32</sub>A-3' and 5'-GATCTA<sub>32</sub>G-3'. *Bgl*II-digested pα19Rβ was ligated to the annealed U<sub>32</sub> sequence to create plasmid pα19Rβ+U32.

The *XbaI*-*SaII* fragment of pα19Rβ+AT×5 (described above) containing (ATTT)<sub>5</sub> was removed, and a *Bgl*II site was created using a synthetic linker, creating plasmid pα19RβΔAU.

***In Vitro* Transcription and Gel Mobility Shift Assay**—Plasmids pα19Rβ+ARE, pα19Rβ+ARE3, pα19Rβ (20), pα19RβΔAU, and pα19Rβ+U32 were linearized by *Bgl*II digestion to yield templates for transcription of ARE, ARE3, Rβ, RβΔAU, and U<sub>32</sub> RNAs, respectively. Plasmids pα19Rβ+AT×1, pα19Rβ+AT×2, pα19Rβ+AT×3, pα19Rβ+AT×4, and pα19Rβ+AT×5 were digested with *SaII* to create the templates for RNA substrates containing AUUUA, (AUUU)<sub>2</sub>, (AUUU)<sub>3</sub>, (AUUU)<sub>4</sub>, and (AUUU)<sub>5</sub>, respectively. Plasmid pGEMmyc (AT1) (20) was linearized with *SspI* for synthesis of *c-myc* ARE. Radiolabeled RNAs were prepared by *in vitro* transcription of each DNA template using either T3 or SP6 RNA Polymerase and [ $\alpha$ -<sup>32</sup>P]UTP (ICN).

Purified recombinant His<sub>6</sub>-AUF1 fusion protein was incubated with 1 fmol of <sup>32</sup>P-labeled RNA probe in a final volume of 10 μl containing 10 mM Tris-HCl (pH 7.5), 5 mM magnesium acetate, 100 mM potassium acetate, 2 mM dithiothreitol, 0.1 mM spermine, 0.1 μg/ml bovine serum albumin, 8 units of RNasin, 0.2 μg/ml tRNA, 5 μg/ml heparin, and 0.1 μg/ml poly(C). Reaction mixtures were incubated on ice for 10 min. Complexes were resolved by electrophoresis through nondenaturing 6% polyacrylamide gels (acrylamide/bisacrylamide ratio of 60:1) in 45 mM Tris borate (pH 8.3), 1 mM EDTA. Gels were prerun for 30 min at 13 V/cm prior to sample loading. Gels were then run at 13 V/cm for 2–3 h, dried, and visualized on a PhosphorImager (Molecular Dynamics). In some experiments there was a loss of shifted products during electrophoresis. This was observed as smears migrating between the bound and free RNA bands. Thus, free RNA bands were routinely used for quantitation using ImageQuant image analysis software (Molecular Dynamics). Free probe concentration was plotted versus His<sub>6</sub>-AUF1



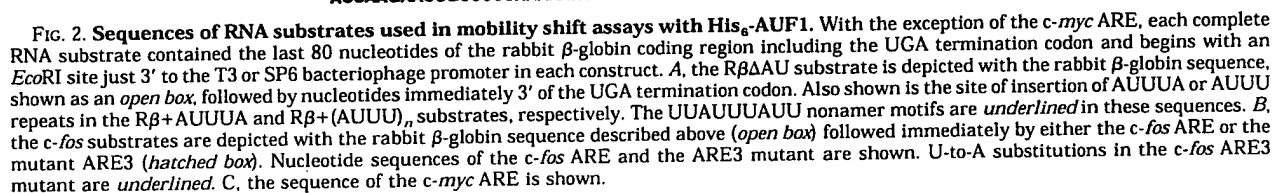
**FIG. 1. Characterization of purified recombinant His<sub>6</sub>-AUF1 protein.** Left panel, Coomassie Blue staining of His<sub>6</sub>-AUF1 protein. One microgram of purified recombinant protein was fractionated in an SDS, 10% polyacrylamide gel. The gel was then stained for protein visualization. The apparent molecular mass of full-length His<sub>6</sub>-AUF1 is 51 kDa (arrow, lane 2). Lane 1 shows prestained molecular mass markers. Right panel, UV cross-linking to *c-fos* ARE. A binding reaction containing 25 ng of purified recombinant protein and 40 fmol of radiolabeled *c-fos* ARE RNA was treated with UV light and digested with RNase A as described under "Materials and Methods." The reaction was fractionated in an SDS, 10% polyacrylamide gel, and the protein bound to RNA was detected by autoradiography. The 51-kDa fusion protein is the major cross-linked species (arrow).

concentration, and apparent  $K_d$  values were determined as the protein concentration at which 50% of the RNA was bound (27). For each RNA substrate tested, binding assays were performed in triplicate, and the average apparent  $K_d$  and standard deviation were determined. In addition, the Newman-Kuels analysis of variance test was applied to each set of apparent  $K_d$  values to identify significant differences ( $p < 0.05$ ) between  $K_d$  values for binding to the various RNA substrates.

**UV Cross-linking**—Twenty-five nanograms of purified recombinant His<sub>6</sub>-AUF1 and 40 fmol of <sup>32</sup>P-labeled RNA containing the *c-fos* ARE were incubated as described above for mobility shift reactions. The binding mixture was then treated with 2500 μJ of UV light for 3 min and digested with 10 μg of RNase A at 37 °C for 30 min. The binding reaction was fractionated on an SDS, 10% polyacrylamide gel using prestained molecular weight standards as markers, fixed in the gel, and detected by autoradiography.

## RESULTS

**Binding of His<sub>6</sub>-AUF1 to Synthetic AREs Containing Tandem Repeats of AUUU**—We previously reported the purification, molecular cloning, and characterization of the ARE-binding protein AUF1. Our hypothesis is that the binding affinity of AUF1 for an ARE should reflect the potency of that ARE as a mRNA destabilizer. To address this hypothesis biochemically, we have examined the RNA-binding properties of AUF1. For this purpose we expressed a His<sub>6</sub>-AUF1 fusion protein in *Escherichia coli* and purified the recombinant protein, which has an apparent molecular mass of about 51 kDa by SDS-polyacrylamide gel electrophoresis (Fig. 1, left panel). To confirm that this polypeptide was the ARE-binding protein, a binding reaction containing radiolabeled *c-fos* ARE was treated with UV light, digested with RNase A, fractionated by SDS-PAGE, and detected by autoradiography. The 51-kDa His<sub>6</sub>-AUF1 polypeptide was the major cross-linked species observed (Fig. 1, right panel). As a control, a lysate prepared from bacteria expressing the empty vector had no detectable ARE-binding activity by UV cross-linking assay (data not shown). Therefore, the RNA-binding protein in these studies is His<sub>6</sub>-AUF1. In addition, the gel filtration profile of His<sub>6</sub>-AUF1 in 10 mM Tris-HCl (pH 7.5), 5 mM magnesium acetate, 100 mM potassium acetate was the same at both 400 nM and 4 μM protein (data not shown), suggesting that the protein does not form high molecular weight aggregates that could affect its ARE-binding activity.



Since multiple tandem repeats of AUUU constitute primarily U-rich sequence, the increasing binding affinity by His<sub>6</sub>-AUF1 observed with increasing AUUU copy number could result from an increase in the number of uridylyate residues. To test His<sub>6</sub>-AUF1 binding affinity for U-rich RNA, binding to a substrate containing U<sub>32</sub> was performed, since this sequence has been tested for mRNA destabilizing activity (17). The apparent  $K_d$  for His<sub>6</sub>-AUF1 binding to the Rβ+U<sub>32</sub> substrate is >500 nM (Fig. 4). While low affinity binding occurs to the Rβ+U<sub>32</sub> substrate, no binding is detected to the control β-globin substrate lacking the U<sub>32</sub> sequence (Fig. 4). Nonetheless the estimated

**Binding of His<sub>6</sub>-AUF1 to Authentic *c-fos* and *c-myc* ARE Sequences**—Our results show that His<sub>6</sub>-AUF1 has a relatively low affinity for a substrate with two tandem AUUU motifs (Table I). The *c-fos* ARE has two tandem AUUU motifs and a single AUUUA motif separated by 19 nucleotides, while the *c-myc* ARE has two AUUUA motifs separated by 25 nucleotides (see Fig. 2). However, the *c-fos* and *c-myc* AREs are very potent mRNA destabilizers (17, 28). We therefore examined binding of His<sub>6</sub>-AUF1 to RNA substrates containing these AREs using the mobility shift assay. A plot of free RNA concentration *versus* fusion protein concentration revealed an apparent  $K_d$  of  $7.8 \pm 0.4$  nM for the *c-fos* ARE (R $\beta$ +fosARE; Fig. 5A). High affinity binding does not require intact AUUUA motifs, since His<sub>6</sub>-AUF1 binds the ARE3 mutant *c-fos* substrate containing single U-to-A substitutions in each AUUUA motif with an apparent  $K_d$  of  $20 \pm 4$  nM (R $\beta$ +ARE3; Fig. 5B). (The difference in binding affinity between the wild-type and mutant *c-fos* ARE is not statistically significant ( $p > 0.05$ ).) Likewise, His<sub>6</sub>-AUF1 binds the *c-myc* ARE with an affinity ( $K_d = 21 \pm 3$  nM) similar to that for the *c-fos* ARE (Table I). (The difference in binding affinity between the *c-myc* and *c-fos* ARE is not statistically significant ( $p > 0.05$ ).) We conclude that AUF1 binds authentic *c-fos* and *c-myc* AREs with high affinity even though these AREs lack

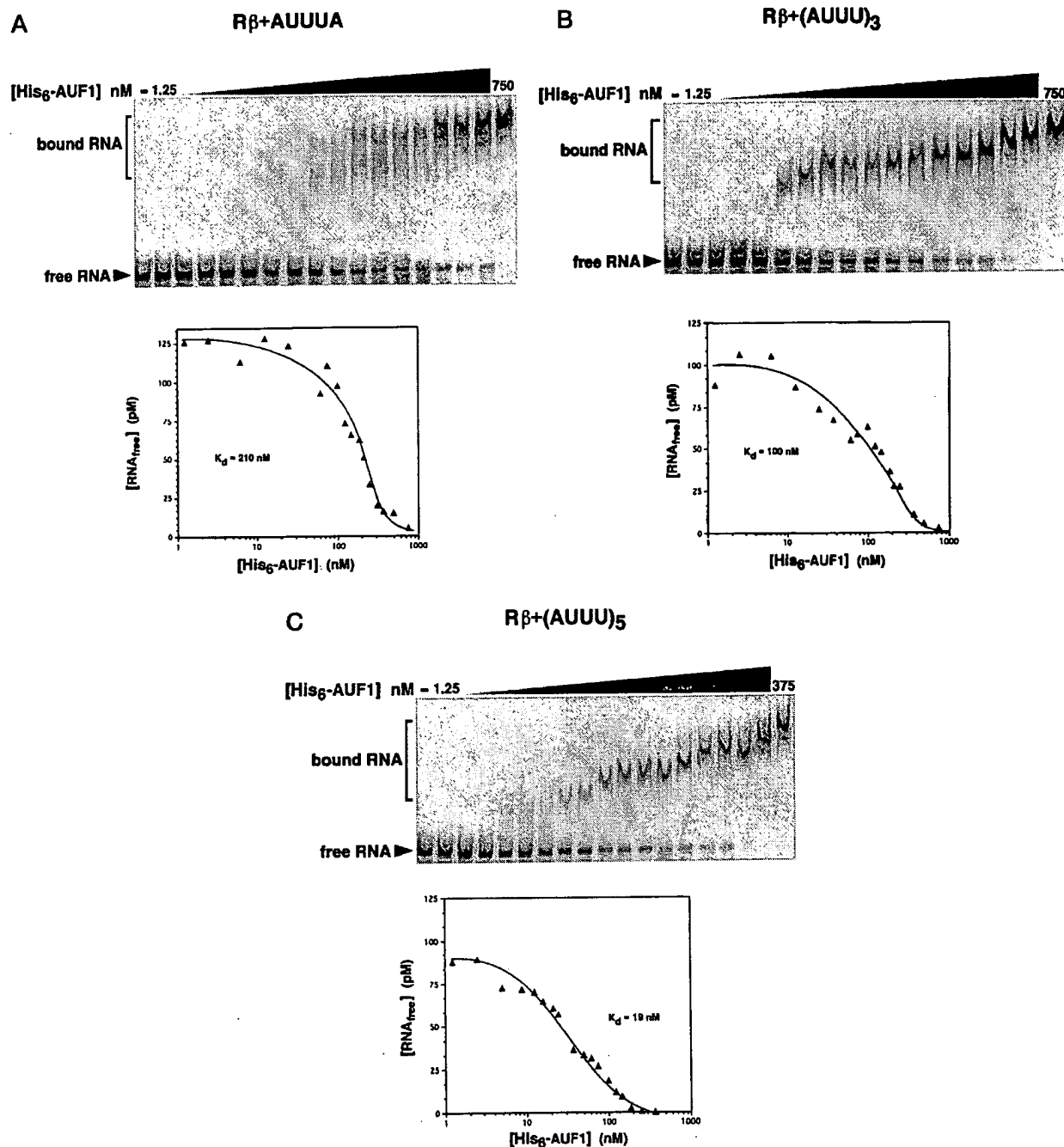


FIG. 3. His<sub>6</sub>-AUF1 binding affinity for AUUU-containing sequences increases as the number of tandemly repeated AUUU motifs is increased. His<sub>6</sub>-AUF1 binding to RNA substrates containing AUUUA or tandem repeats of AUUU (see Fig. 2) was analyzed by electrophoretic mobility shift assays and apparent  $K_d$  for His<sub>6</sub>-AUF1 binding to each RNA was determined as described under "Materials and Methods." Representative binding reactions using His<sub>6</sub>-AUF1 and radiolabeled RNA substrates are shown at the top of each panel, and plots of [RNA<sub>free</sub>] versus [His<sub>6</sub>-AUF1] are shown at the bottom of each panel. A, Rβ+AUUUA; B, Rβ+(AUUU)<sub>3</sub>; C, Rβ+(AUUU)<sub>5</sub>. For each RNA substrate, the  $K_d$  value shown is the average of three separate experiments.

multiple tandem repeats of AUUU. Thus, AUF1 is capable of binding a number of different AREs with high affinity.

#### DISCUSSION

The importance of AREs for mRNA turnover was first realized in 1986 (12, 14), yet it is still unclear how AREs function in mRNA decay. We have utilized biochemical approaches to identify *trans*-acting factors that bind AREs in order to relate such RNA binding to mRNA degradation. We previously purified, characterized, and molecularly cloned the ARE-binding

protein AUF1, and in the present study we have examined binding affinities of a His<sub>6</sub>-AUF1 fusion protein for A+U-rich sequences with defined relative potencies as mRNA destabilizers. Here, by determining apparent  $K_d$  values for His<sub>6</sub>-AUF1 binding, we demonstrate that AUF1-ARE binding affinity is directly related to the potency with which an ARE destabilizes a heterologous mRNA (Table I). Additionally, the affinity of His<sub>6</sub>-AUF1 for the most potent destabilizing AREs is within the average range ( $10^{-9}$  M) of affinities exhibited by several other RNA-binding proteins that recognize specific sequences or

TABLE I  
Summary of apparent  $K_d$  values for His<sub>6</sub>-AUF1 binding to RNA substrates with corresponding degradation rates of the heterologous mRNAs

The binding substrates in electrophoretic mobility shift assays are represented in the left column. Relative degradation rates, summarized from previous studies (Refs. 16 and 17), are depicted in the middle column as a series of plus symbols, where the number of symbols is proportional to the degradation rate of the mRNA (one is slowest and four is fastest). The apparent  $K_d$  values (average  $\pm$  S.E.) for His<sub>6</sub>-AUF1 binding to each RNA are shown in the right column.

RNA	Relative degradation rate	Apparent $K_d$ for AUF1 binding
		<i>nM</i>
<i>c-fos</i> ARE	++++	$7.8 \pm 0.4$
<i>c-myc</i> ARE	++++	$21 \pm 3$
(AUUU) <sub>5</sub>	++++	$19 \pm 7$
(AUUU) <sub>4</sub>	+++	$60 \pm 20$
(AUUU) <sub>3</sub>	++	$100 \pm 50$
(AUUU) <sub>2</sub>	+	$150 \pm 30$
AUUUA	+	$210 \pm 50$
U <sub>32</sub>	+	>500
R $\beta$	+	>750

structures (29).

Certain A+U-rich sequences are more potent mRNA destabilizers than others, suggesting that the potencies of destabilizers are proportional to the binding affinities of a cellular decay factor(s). For example, when placed in the context of a heterologous, normally stable mRNA, AUUUA and (AUUU)<sub>2</sub> are relatively ineffective as destabilizing elements; (AUUU)<sub>3</sub> has a modest destabilizing effect; (AUUU)<sub>4</sub> increases the decay rate further; and (AUUU)<sub>5</sub> is the most potent destabilizer of the five (16, 17, 26). In fact, (AUUU)<sub>5</sub> increases the degradation rate of a reporter mRNA to about the same extent as does the *c-fos* ARE (17). Likewise, His<sub>6</sub>-AUF1 binds the *c-fos* ARE ( $K_d = 7.8 \pm 0.4$  nM) and the (AUUU)<sub>5</sub> substrate ( $K_d = 19 \pm 7$  nM) with similar affinities. (The differences are not statistically significant ( $p > 0.05$ ).) Statistical analyses were used to determine significant differences (*i.e.*  $p < 0.05$ ) between  $K_d$  values for His<sub>6</sub>-AUF1 binding to various RNA substrates, and as a result the RNA sequences used in this study can be grouped into three general classes: (i) RNAs that are either not bound or bound with low affinity by AUF1 and are not mRNA destabilizers ( $\beta$ -globin, U<sub>32</sub>, and AUUUA); (ii) RNAs that are bound with gradually increasing, moderate affinities by AUF1 and have a gradually increasing, partial destabilizing effect ((AUUU)<sub>2</sub> < (AUUU)<sub>3</sub> < (AUUU)<sub>4</sub>); and (iii) RNAs that are bound with the highest affinity by AUF1 and are potent mRNA destabilizers (*c-fos* and *c-myc* AREs and (AUUU)<sub>5</sub>). Based upon these ranges of AUF1 binding affinities for various RNA substrates (low, moderate, high) and the relationship of high affinity binding to mRNA decay, the affinity of AUF1 for a mRNA may dictate the rate at which it is degraded. Therefore, cellular AUF1 concentration may be one determinant of mRNA half-life. In this regard we found that by comparing Western blots of K562 cytoplasmic extracts with known amounts of purified recombinant p37<sup>AUF1</sup> (the isoform used in these studies) that there are approximately  $3.2 \times 10^4$  cytoplasmic molecules of p37<sup>AUF1</sup>/cell (data not shown). Assuming a diameter of 20  $\mu$ m for K562 cells and 50% of the cell volume as cytoplasm (30), the concentration of p37<sup>AUF1</sup> is approximately 25 nM. This value is comparable with the apparent  $K_d$  for binding to the *c-myc* ARE. Thus low cellular concentrations of active AUF1 may be sufficient for binding to a mRNA that contains a high affinity AUF1-binding site such as the *c-myc* and *c-fos* AREs. Based upon our results, such mRNAs should have very short half-lives. Likewise, mRNAs with AREs bound with lower affinities by AUF1 might require a higher concentration of active AUF1 for binding;

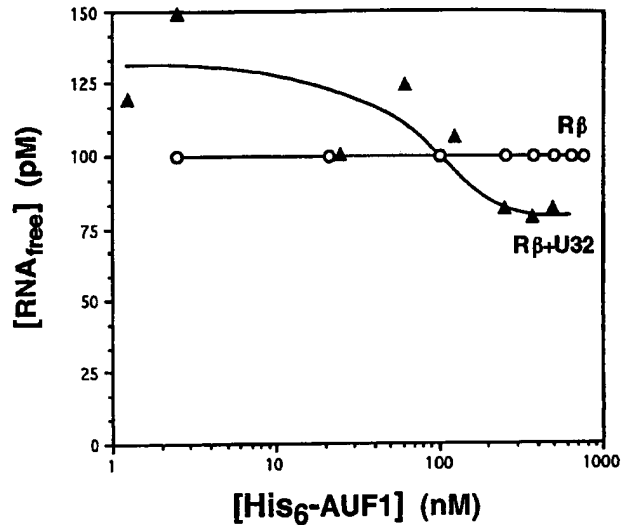


FIG. 4. R $\beta$ +U<sub>32</sub> is not a high affinity His<sub>6</sub>-AUF1 binding substrate. RNA substrates contained the last 80 nucleotides of the rabbit  $\beta$ -globin coding region alone (R $\beta$ ) or linked to U<sub>32</sub> (R $\beta$ +U<sub>32</sub>). Binding affinity of His<sub>6</sub>-AUF1 for the R $\beta$ +U<sub>32</sub> or R $\beta$  substrates was determined by electrophoretic mobility shift assays as described under "Materials and Methods." Representative plots of [RNA<sub>free</sub>] versus [His<sub>6</sub>-AUF1] for R $\beta$ +U<sub>32</sub> (triangles) and R $\beta$  (open circles) are shown. The apparent  $K_d$  for binding to R $\beta$ +U<sub>32</sub> was determined from three separate experiments to be >500 nM, which was the highest protein concentration used. No binding to R $\beta$  was detected.

these mRNAs should be degraded at a slower rate than those with high affinity binding sites. Thus, the availability of active AUF1 for ARE binding is a potential mechanism by which cells could control mRNA turnover rates and one in which the decay of multiple mRNAs could be differentially regulated by AUF1 concentration. Support for this hypothesis is the relationship between AUF1 levels and ARE-directed mRNA destabilization observed in DDT1-MF2 hamster smooth muscle cells treated with (-)-isoproterenol. In this case, (-)-isoproterenol induces an increase in cellular AUF1 protein and mRNA levels. This increase in turn correlates with a faster decay rate for  $\beta_2$ -adrenergic receptor mRNA, which contains an AUF1 binding site(s) in the 3'-UTR (25).

Recently, two groups reported that the functional sequence within an ARE appears to be the nonamer sequence UUAUUUAUU (16, 17). One copy of the sequence UUAUUUAUU in the 3'-UTR of a normally stable mRNA can increase its degradation rate compared with the wild-type mRNA, while two copies of the nonamer motif act as a very potent mRNA destabilizer. As depicted in Fig. 2, (AUUU)<sub>3</sub> contains one copy of the nonamer (underlined in Fig. 2); (AUUU)<sub>4</sub> contains two overlapping copies; and (AUUU)<sub>5</sub> contains two copies that overlap by a single nucleotide. Consistent with the potencies of two copies or one copy of the nonamer as destabilizers, binding affinity of His<sub>6</sub>-AUF1 for (AUUU)<sub>5</sub> is 3-fold and 5-fold greater than the binding affinities for (AUUU)<sub>4</sub> or (AUUU)<sub>3</sub>, respectively (Table I). Thus, AUF1 may function in part via recognition of the nonamer motif.

Despite the potential importance of the nonamer motif UUAUUUAUU in ARE-directed mRNA decay, it is important to note that not all AREs found in unstable mRNAs contain this motif. For example, the portion of the *c-myc* ARE that functions as a very potent mRNA destabilizer does not contain this motif (28). While the *c-myc* ARE does contain noncontiguous AUUUA motifs, destabilizing AREs that contain no AUUUA motifs have also been identified. In addition, the presence of one or more AUUUA motifs in an ARE may not be sufficient for effective

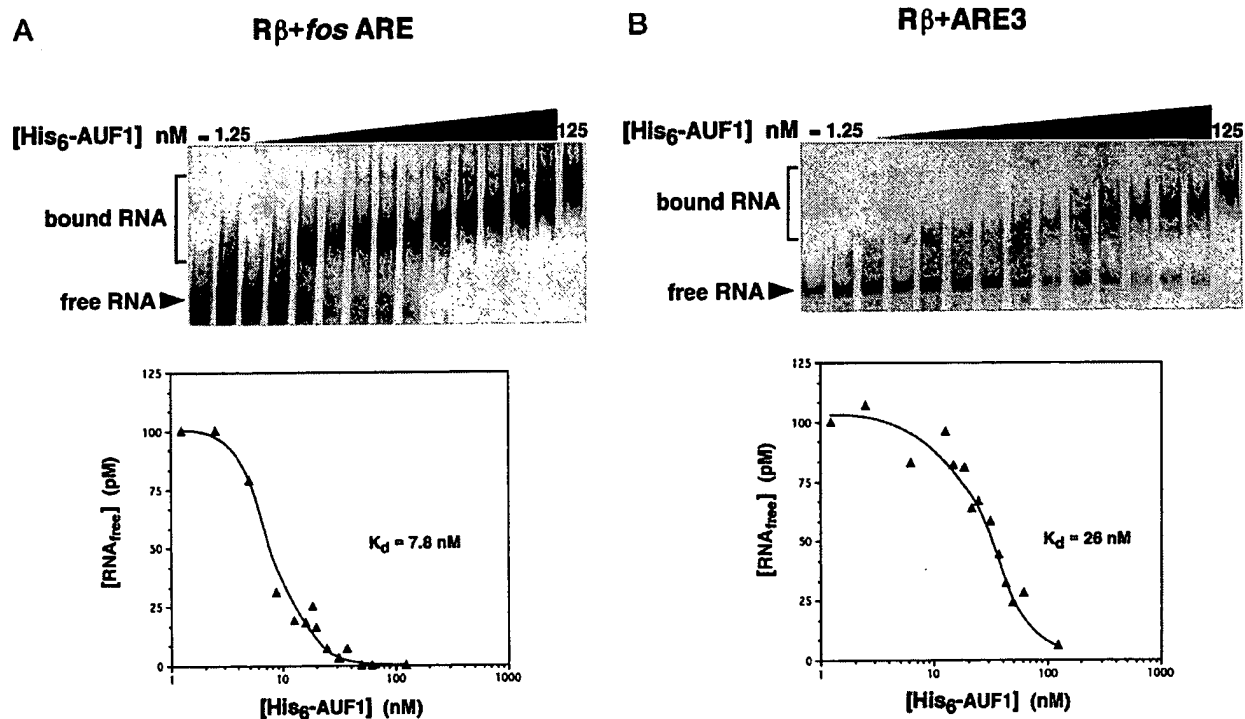


FIG. 5. High affinity binding of His<sub>6</sub>-AUF1 to *c-fos* ARE substrates. His<sub>6</sub>-AUF1 binding to RNA substrates containing the wild-type *c-fos* ARE or the ARE3 mutant (see Fig. 2) was analyzed by electrophoretic mobility shift assays, and apparent  $K_d$  values for His<sub>6</sub>-AUF1 binding to each RNA were determined as described under "Materials and Methods." Representative binding reactions using His<sub>6</sub>-AUF1 and radiolabeled RNA substrates are shown at the top of each panel, and plots of [RNA<sub>free</sub>] versus [His<sub>6</sub>-AUF1] are shown at the bottom of each panel. A, Rβ+*fos* ARE; B, Rβ+*fos* ARE3. For each RNA substrate, the  $K_d$  value shown is the average of three separate experiments.

mRNA destabilization (28). Moreover, analysis of a *c-fos* ARE mutant, ARE3, with single U-to-A substitutions in all three AUUUA motifs showed that intact AUUUA motifs are not required for rapid mRNA deadenylation but are important for rapid degradation of the mRNA body (9). The apparent  $K_d$  for His<sub>6</sub>-AUF1 binding to the *c-fos* ARE and the ARE3 mutant were not statistically different ( $7.8 \pm 0.4$  nM and  $20 \pm 4$  nM, respectively;  $p > 0.05$ ). Likewise, His<sub>6</sub>-AUF1 bound a mutant *c-myc* ARE with single U-to-A mutations in both AUUUA motifs with affinity similar to wild-type *c-myc* ARE (data not shown). Thus, intact AUUUA motifs are not required for high affinity binding of His<sub>6</sub>-AUF1 to the *c-fos* and *c-myc* AREs. Therefore, although the study of AUUUA and UUAUUUU motifs has contributed greatly to understanding of ARE-directed mRNA decay, it is evident that these motifs may constitute only a subset of important motifs within various AREs.

In conclusion, our results suggest that the affinity of AUF1 for particular ARE sequences is related to their potency as mRNA destabilizers. Future experiments will utilize AUF1 as a tool to define multiple classes of AREs and to define specific nucleotide requirements for AUF1 binding by selection of high affinity binding substrates from combinatorial libraries of RNA sequences (e.g., SELEX; reviewed in Ref. 31).

**Acknowledgments**—We thank Gray Shaw for the pNEORβG<sup>AT</sup> plasmids, Ann-Bin Shyu for the *c-fos* plasmids, Jim Rose for help with statistical analyses, and Paul Bohjanen, Mariano Garcia-Blanco, Doug Lyles, David Ornelles, and Jeff Ross for comments on the manuscript.

#### REFERENCES

- Beelman, C. A., and Parker, R. (1995) *Cell* 81, 179–183
- Peltz, S. W., Brewer, G., Bernstein, P., Hart, P. A., and Ross, J. (1991) *Crit. Rev. Eukaryotic Gene Expr.* 1, 99–126
- Sachs, A. (1993) *Cell* 74, 413–421
- Carter, B. Z., and Malter, J. S. (1991) *Lab. Invest.* 65, 610–621
- Schiavi, S. C., Belasco, J. G., and Greenberg, M. E. (1992) *Biochim. Biophys. Acta* 1114, 95–106
- Ross, J. (1995) *Microbiol. Rev.* 59, 423–450
- Greenberg, M. E., and Belasco, J. (1993) *Control of Messenger RNA Stability* (Brawerman, G., and Belasco, J., eds) pp. 199–218. Academic Press, Inc., San Diego, CA
- Brewer, G., and Ross, J. (1988) *Mol. Cell. Biol.* 8, 1697–1708
- Shyu, A.-B., Belasco, J. G., and Greenberg, M. E. (1991) *Genes & Dev.* 5, 221–231
- Swartwout, S. G., and Kinniburgh, A. J. (1989) *Mol. Cell. Biol.* 9, 288–295
- Bernstein, P., and Ross, J. (1989) *Trends Biochem. Sci.* 14, 373–377
- Caput, D., Beutler, B., Hartog, K., Thayer, R., Brown-Shimer, S., and Cerami, A. (1986) *Proc. Natl. Acad. Sci. U. S. A.* 83, 1670–1674
- Jones, T. R., and Cole, M. D. (1987) *Mol. Cell. Biol.* 7, 4513–4521
- Shaw, G., and Kamen, R. (1986) *Cell* 46, 659–667
- Wilson, T., and Treisman, R. (1988) *Nature* 336, 396–399
- Lagnado, C. A., Brown, C. Y., and Goodall, G. J. (1994) *Mol. Cell. Biol.* 14, 7984–7995
- Zubiaga, A. M., Belasco, J. G., and Greenberg, M. E. (1995) *Mol. Cell. Biol.* 15, 2219–2230
- Brewer, G., and Ross, J. (1989) *Mol. Cell. Biol.* 9, 1996–2006
- Brewer, G. (1991) *Mol. Cell. Biol.* 11, 2460–2466
- Zhang, W., Wagner, B. J., Ehrenman, K., Schaefer, A. W., DeMaria, C. T., Crater, D., DeHaven, K., Long, L., and Brewer, G. (1993) *Mol. Cell. Biol.* 13, 7652–7665
- Burd, C. G., and Dreyfuss, G. (1994) *Science* 265, 615–621
- Ehrenman, K., Long, L., Wagner, B. J., and Brewer, G. (1994) *Gene (Amst.)* 149, 315–319
- Hadcock, J. R., Wang, H., and Malbon, C. C. (1989) *J. Biol. Chem.* 264, 19928–19933
- Tremmel, K. D., Pende, A., Bristow, M. R., and Port, J. D. (1996) *Keystone Symposia: The Molecular Biology of the Cardiovascular System*, p. 47. Keystone Symposia, Silverthorn, CO
- Pende, A., Tremmel, K. D., DeMaria, C. T., Blaxall, B. C., Minobe, W. A., Sherman, J. S., Bisognano, J. D., Bristow, M. R., Brewer, G., and Port, J. D. (1996) *J. Biol. Chem.* 271, 8493–8501
- Akashi, M., Shaw, G., Hachiya, M., Elstner, E., Suzuki, G., and Koeffler, P. (1994) *Blood* 83, 3182–3187
- Carey, J. (1991) *Methods Enzymol.* 208, 103–117
- Chen, C.-Y. A., and Shyu, A.-B. S. (1994) *Mol. Cell. Biol.* 14, 8471–8482
- McCarthy, J. E. G., and Kolmus, H. (1995) *Trends Biochem. Sci.* 20, 191–197
- Koeffler, H. P., and Golde, D. W. (1980) *Blood* 56, 344–350
- Kenan, D. J., Tsai, D. E., and Keene, J. D. (1994) *Trends Biochem. Sci.* 19, 57–64

## The Neuronal RNA Binding Protein Nova-1 Recognizes Specific RNA Targets In Vitro and In Vivo

RONALD J. BUCKANOVICH AND ROBERT B. DARNELL\*

Laboratory of Molecular Neuro-Oncology, The Rockefeller University, New York, New York 10021

Received 7 January 1997/Returned for modification 24 February 1997/Accepted 11 March 1997

**Nova-1, an autoantigen in paraneoplastic opsoclonus myoclonus ataxia (POMA), a disorder associated with breast cancer and motor dysfunction, is a neuron-specific nuclear RNA binding protein. We have identified in vivo Nova-1 RNA ligands by combining affinity-elution-based RNA selection with protein-RNA immunoprecipitation. Starting with a pool of  $\sim 10^{15}$  random 52-mer RNAs, we identified long stem-loop RNA ligands that bind to Nova-1 with high affinity ( $K_d$  of  $\sim 2$  nM). The loop region of these RNAs harbors a  $\sim 15$ -bp pyrimidine-rich element [UCAU(N)<sub>0-2</sub>]<sub>3</sub> which is essential for Nova-1 binding. Mutagenesis studies defined the third KH domain of Nova-1 and the [UCAU(N)<sub>0-2</sub>]<sub>3</sub> element as necessary for in vitro binding. Consensus [UCAU(N)<sub>0-2</sub>]<sub>3</sub> elements were identified in two neuronal pre-mRNAs, one encoding the inhibitory glycine receptor  $\alpha 2$  (GlyR  $\alpha 2$ ) and a second encoding Nova-1 itself. Nova-1 protein binds these RNAs with high affinity and specificity in vitro, and this binding can be blocked by POMA antisera. Moreover, both Nova-1 and GlyR  $\alpha 2$  pre-mRNAs specifically coimmunoprecipitated with Nova-1 protein from brain extracts. Thus, Nova-1 functions as a sequence-specific nuclear RNA binding protein in vivo; disruption of the specific interaction between Nova-1 and GlyR  $\alpha 2$  pre-mRNA may underlie the motor dysfunction seen in POMA.**

RNA-protein interactions are important in the posttranscriptional regulation of RNA metabolism and expression. Nascent RNA transcripts associate with large multiprotein complexes that include hnRNP proteins and snRNP particles (19). RNA protein complexes subsequently participate in polyadenylation, RNA splicing, RNA transport, and translational control. Defining target RNAs with which RNA binding proteins (RBPs) interact has been critical in defining their function. In *Drosophila melanogaster*, the identification of sequence-specific targets for the sex-lethal (sxl) (5, 30) and transformer-2 (28) RBPs has led to a precise understanding of their role in regulating RNA splicing. In mammals, the ability of U2AF(65) to bind to polypyrimidine tracts is believed to recruit U2 snRNP to the branch site (48, 69). Identification of specific RNA ligands for the human immunodeficiency virus Rev and Tat proteins has clarified their role in regulating viral RNA transcription, processing, and transport (14, 27, 34, 40, 70). There remain, however, many RBPs for which specific RNA targets have not been identified.

RNA selection has been used as an in vitro approach to identify RNA ligands for RBPs (61, 65). This approach was originally used to confirm the specificity of rRNA binding sites recognized by T4 DNA polymerase (65) and has subsequently been used to confirm and extend known binding sites for the U1-snRNP-A protein (64), sxl (55), and the viral Rev (3, 31) and Tat (66) proteins. RNA selection experiments have also been used in efforts to identify RNA ligands for RBPs that do not have known sequence-specific binding sites, including hnRNP proteins (12), SR proteins (29, 62), and the neuronal RBP (n-RBP) Hel-N1 (36), although in general these studies have yielded short in vitro consensus RNA ligands whose in vivo significance is currently being explored.

Nova-1 is a nuclear n-RBP identified as a target antigen in paraneoplastic opsoclonus myoclonus ataxia (POMA) (9, 10,

16). POMA is a neurologic disorder that develops when systemic tumors (typically breast tumors) ectopically express Nova, triggering the production of an autoimmune response characterized by high-titer anti-Nova antibodies in the cerebrospinal fluid (9, 38). The neurologic symptoms of POMA are thought to result, at least in part, from the failure of inhibition of brainstem and/or spinal motor systems (16). Cloning of the Nova-1 gene led to the demonstration that it was an n-RBP composed of three KH-type RNA binding domains (9, 10). Nova-1, together with the Hu family of n-RBPs, are distinct from most mammalian RBPs in that their expression is extremely tissue specific. For example, most hnRNP proteins are ubiquitously expressed, although some show different levels of expression among different tissues (13, 33). Nova-1 is expressed only in neurons of the subcortical central nervous system (9, 10).

KH-type RBPs are an expanding family of proteins for which some functional data are available, although whether they act as sequence-specific RBPs is uncertain. KH-type proteins include FMR-1 (encoded by the fragile-X gene), which harbors two KH domains; loss of function mutations of the fragile-X gene is associated with mental retardation (11, 58). Nova-1 is also related to KH-type RBPs involved in regulating RNA splicing, including MER-1 in yeast (21), PSI in *Drosophila* (54), and the recently described mammalian proteins KSR and SF-1 (2, 41). Nova-1 is most closely related to hnRNP K, a protein whose function or targets as an RBP are unknown. In this paper, we report the identification of a specific stem-loop RNA which Nova-1 binds in vitro and in vivo in two neuronal pre-mRNAs. We discuss these findings in relation to the biology of Nova as an RBP and as a disease antigen.

### MATERIALS AND METHODS

**RNA selection.** An oligonucleotide harboring a 52-bp random sequence surrounded by primer binding sites (GGG ACA ATT CCG ACC AGA AG N<sub>52</sub> TAT GTG CGT CTA CAT GGA TCC TCA [22]) was synthesized on an ABI DNA synthesizer to yield 2 mg of DNA with an estimated complexity of  $\sim 3 \times 10^{16}$  sequences; the oligonucleotide was characterized and PCR amplified by using forward and reverse oligonucleotide primers as described previously (20, 22). Following PCR amplification, the sequences of 24 random clones from this

\* Corresponding author. Mailing address: Laboratory of Molecular Neuro-Oncology, The Rockefeller University, 1230 York Ave., New York, NY 10021. Phone: (212) 327-7460. Fax: (212) 327-7147. E-mail: darnell@rockvax.rockefeller.edu.

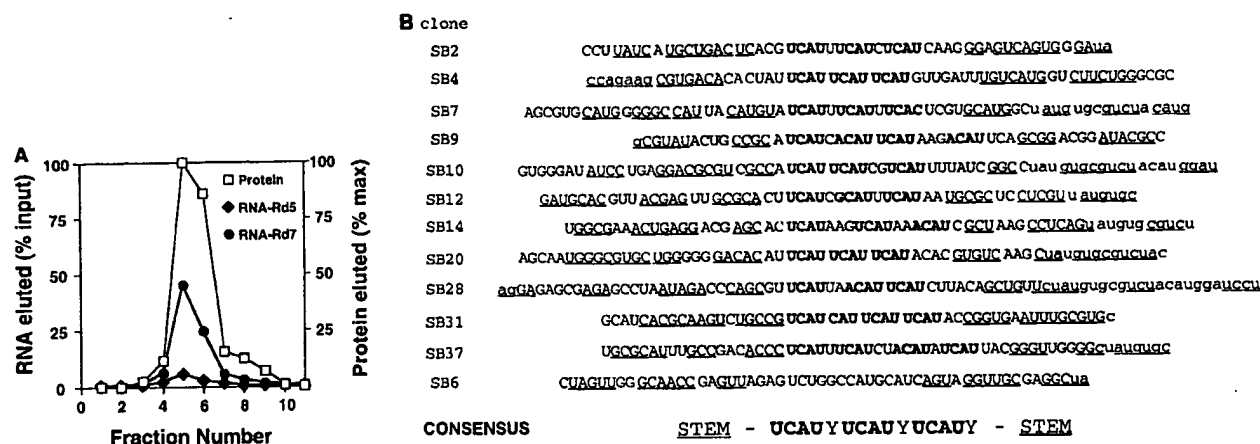


FIG. 1. Selection of Nova-1 RNA ligands from a random RNA library. (A) Bound RNAs were enriched in each round of selection by specifically eluting histidine-tagged Nova-1 fusion protein in 1.0 M imidazole. A typical protein elution curve is indicated (open squares), with each point defined as a percentage of protein eluted in the peak fraction. RNA elutions from selection round 5 (RNA-Rd5) or round 7 (RNA-Rd7) are shown. Multiple selection cycles yielded RNA pools in which ~87% of input RNA counts coeluted with Nova-1 protein (RNA selection round 7; see also Table 1). (B) Sequences of 12 of 14 PCR clones encoding RNA ligands cloned after seven rounds of RNA selection are shown; two additional clones in which no consensus was identified are not shown. UCAU repeats in each ligand are shown in boldface. Regions of the RNA which form the inverted repeat stem are underlined. Stem sequences which are derived from the nonrandom regions of the RNA ligand (primer binding sites) are in lowercase. A consensus sequence is shown below.

pool were determined; each clone was unique, and the overall base composition showed a slight excess of purines (24% A and 33% G) over pyrimidines (21% C and 20% U) and a random distribution of dinucleotide frequencies (data not shown). Two pool equivalents of purified library DNA (110 µg,  $\sim 2.1 \times 10^{14}$  molecules) were transcribed *in vitro* with T7 RNA polymerase and [ $\alpha$ - $^{32}$ P]UTP as described previously (20). RNA was purified from denaturing acrylamide gels, heated to 70°C for 5 min, and applied to a precolumn containing an irrelevant histidine-tagged fusion protein to adsorb nonspecifically bound RNAs. The flowthrough was then applied to a histidine-tagged Nova-1 nickel affinity column in high-salt buffer (HSB; 0.5 M LiCl, 20 mM Tris-HCl [pH 7.6], 1 mM MgCl<sub>2</sub>, 60 mM imidazole [20, 25]). Following washing in 10 column volumes of HSB, protein and RNA were coeluted by the addition of 1.0 M imidazole in HSB. Protein-bound RNA was extracted in phenol-0.5% sodium dodecyl sulfate at 50°C and ethanol precipitated with LiCl and glycogen.

**Filter binding assays.** Filter binding using purified histidine-tagged Nova fusion proteins and methods previously described (10). Briefly, the indicated concentrations of fusion proteins were incubated with 50 to 200 fmol of  $^{32}$ P-labeled RNA as indicated; RNA was transcribed directly from PCR products encoding RNA selection clones or from oligonucleotides harboring T7 RNA polymerase promoters. Binding was performed in 100 µl of binding buffer (0.5 M LiCl, 20 mM Tris-HCl [pH 7.4], 1 mM MgCl<sub>2</sub>, 10 µg of yeast tRNA/ml) for 30 min, samples were filtered through nitrocellulose and washed with 1 volume of binding buffer, and retained counts were determined in a scintillation counter.

**Gel shift assays.** Gel shift assays were performed with 50 ng or otherwise indicated amounts of fusion protein in 10 µl of protein buffer (50 µg of bovine serum albumin/ml, 0.5 M LiCl, 20 mM Tris-HCl [pH 7.6], 1 mM MgCl<sub>2</sub>; 50 fmol of  $^{32}$ P-labeled RNA was heated to 80°C for 5 min, cooled to room temperature, and then added to RNA buffer (0.5 M LiCl, 20 mM Tris-HCl [pH 7.6], 1 mM MgCl<sub>2</sub>, 40 ng of yeast tRNA/ml) in a final volume of 10 µl. Protein and RNA samples were mixed and allowed to equilibrate for 30 min at room temperature. For supershift experiments, antibody was then added and incubated an additional 30 min at room temperature. Two percent of each sample was resolved with a 50 mM Tris-glycine-4% polyacrylamide gel, dried, and exposed to XAR-5 autoradiography film.

**Immunoprecipitation.** Adult mouse brains were homogenized with a Dounce homogenizer in NET-Triton (150 mM NaCl, 50 mM Tris-HCl [pH 7.4], 0.1% Triton-X 100, protease inhibitors). Nuclei were collected by centrifugation and resuspended in 1 ml of NET-Triton, sonicated, and spun in a microcentrifuge, and the supernatant was precleared with protein A-Sepharose for 30 min at 4°C, followed by incubation with the indicated antibodies and protein A-Sepharose. After five washes with cold RIPA buffer (150 mM NaCl, 50 mM Tris [pH 7.4], 0.1% sodium dodecyl sulfate, 1% Nonidet P-40, 0.5% deoxycholate), samples were resuspended in DNase buffer (50 mM Tris-HCl [pH 7.4], 6 mM MgCl<sub>2</sub>, 3 mM CaCl<sub>2</sub>) with 20 U of DNase and 20 U of RNase inhibitor at 37°C for 1 h, extracted with phenol-chloroform at 55°C, extracted with chloroform, and precipitated with ethanol and sodium acetate. Reverse transcription (RT)-PCR was performed as previously described (9). All PCRs except the PCR to detect clathrin were performed by using 30 cycles (1 min at 94°C, 30 s at 58°C, and 45 s at 72°C). PCR to detect clathrin DNA was performed by using 32 cycles of 1 min at 94°C, 30 s at 54°C, and 45 s at 72°C. Primers used were as follows: glycine

receptor  $\alpha 2$  (GlyR  $\alpha 2$ ), AAAATACTAGTGGGAAGTTATCATGCA and CAT GGTGGTTTCTGTGACTGATC; Nova-1, GCGAATCTCCAGATCGCATC AAACAA and ACTGAAGGCTCCAAAAGTCTTC; clathrin, TTAACCCCTGT GCTGCTGTCTTG and GCTGCTGGTAGAACCTTTGTGTCAG; HuD, GCACATGAATTACTTGCCAT and CATAGGCCATATTAAGCA; Hn1, ATTGCTGTAAACCAATCTA and ATTCCATCGATCCTCATT; and brain-specific Na<sup>+</sup> channel, GGAATCTGGAAGCTGGATTTCAGTG and TCGGGAATTATCATGGACC.

## RESULTS

We have examined the function of Nova-1 by using an affinity elution-based RNA selection (25, 61) to identify specific Nova-1 protein-RNA interactions. RNA was transcribed from two pool equivalents of a PCR-amplified oligonucleotide library comprised of 52-nucleotide random-mers. The transcribed RNA, which had an estimated complexity of  $2.1 \times 10^{14}$  molecules, was applied to a nickel column containing an irrelevant histidine-tagged fusion protein, and the flowthrough was then applied to a nickel column containing an Nova-1 histidine-tagged fusion protein. After washing, Nova-1 fusion protein, together with bound RNA, was affinity eluted in 1.0 M imidazole. In this way, we were able to monitor the coordinate elution of protein and RNA (Fig. 1A) and found that the fraction of RNA bound to Nova-1 rose from essentially undetectable in the initial pool (<0.05%) to ~87% after seven cycles of selection (Table 1).

Sequencing of individual PCR clones obtained after the seventh round of RNA selection identified a clear RNA consensus sequence (Fig. 1B). There were two components to the consensus, a structural element (an inverted stem sequence) which bounded a sequence-specific element (a loop sequence), which together accounted for nearly the full 52 randomized nucleotides present in each clone of the library RNA. The stem consisted of an average of 13 bp in each arm; there was no apparent sequence similarity between the stem of any two clones. The loop element contained a conserved sequence motif that was nearly identical in each of the 11 clones and consisted of a pyrimidine-rich consensus sequence of a core of three UCAU repeats. Six of eleven clones matching the consensus sequence contain three UCAUs, and four of the five remaining clones have two UCAUs and a third NCAU. The

TABLE 1. Multiple selection cycles yield Nova-1 RNA ligand<sup>a</sup>

Selection round	Nova-1-bound RNA (% of input)
1	≤0.05
2	ND
3	0.26
4	3.4
5	10.2
6	51.0
7	87

<sup>a</sup> Percentages of total RNA input which specifically eluted with Nova-1 protein for each round of RNA selection are shown. The fraction of RNA bound to Nova-1 in each cycle of selection was calculated by dividing the number of counts eluted with Nova-1 protein by the number of counts loaded onto the Nova-1 column. In round 1, RNA was eluted in batch from the Nova column by phenol extraction. ND, not determined.

UCAU repeats were separated by zero to two nucleotides; while these sequences were preferentially (~75%) pyrimidines, this frequency was not statistically significant, and evidence presented below suggests that pyrimidines between the UCAU repeats are not essential for Nova-1 binding. Therefore we have designated the Nova-1 consensus RNA ligand [UCAU(N)<sub>0-2</sub>]<sub>3</sub>.

To characterize the Nova-1-RNA interaction, and to map the domains of Nova-1 which are responsible for recognition of the [UCAU(N)<sub>0-2</sub>]<sub>3</sub> element, we quantitated binding of a selected RNA clone (SB2) to full-length Nova-1 fusion protein (NFP) or to Nova-1 deletion constructs containing either the first (N1), first and second (N1-2), or third (N3) KH domains. In a filter binding assay, NFP demonstrated high-affinity saturable binding to SB2 ( $K_d$  of ~20 nM in 0.5 M LiCl and of  $K_d$  of ~2 nM in 0.1 M LiCl [Fig. 2A and data not shown]). The affinity of N3 for SB2 was reduced approximately 10-fold ( $K_d$  of ~180 nM in 0.5 M LiCl and  $K_d$  of ~10 nM in 0.1 M LiCl), while neither N1, N1-2, nor an N3 protein carrying a single (leucine-to-asparagine) point mutation (N3 L213N [10]) showed detectable binding to SB2 (Fig. 2A). Thus, Nova-1 binds to SB2 with high affinity, and this interaction is mediated at least in part by KH3.

We analyzed the sequence specificity and the structural requirements for Nova-1 binding to the [UCAU(N)<sub>0-2</sub>]<sub>3</sub> RNA ligand by quantitating its interaction with a series of SB2-derived mutant RNA ligands. While NFP showed high-affinity binding to SB2, there was no binding to an RNA (SB2A3U) in which the (UCAU)<sub>3</sub> element was mutated to (UCUU)<sub>3</sub> (Fig. 2B). In addition, NFP binding to SB2 was competed with an excess of cold SB2 RNA (50% inhibition with a ~12-fold molar excess of competitor) but could not be competed with an excess of SB2A3U RNA (see Fig. 4C). A structural mutant, in which base pairing in the stem sequence was eliminated while the [UCAU(N)<sub>0-2</sub>]<sub>3</sub> sequences were unchanged, showed a threefold reduction in binding of NFP (Fig. 2B). These results were confirmed by gel shift assays. Figure 2C shows that the SB2 RNA was almost completely shifted after binding to NFP but unaffected following incubation with an equimolar amount of an irrelevant control protein (amino acids 270 to 413 of cdr2 [16, 16a]). In contrast, the mutant SB2A3U RNA showed no gel shift when bound to NFP (Fig. 2C). We also found that the SB2-NFP complexes could be supershifted by using Nova-1 antibodies (Fig. 2C). SB2-NFP complexes were not shifted when incubated with preimmune antiserum or an irrelevant affinity-purified antibody, and anti-Nova-1 antibodies alone did not shift SB2 RNA (Fig. 2C and data not shown). We conclude that Nova-1 is a sequence-specific RBP. The preferred Nova-1

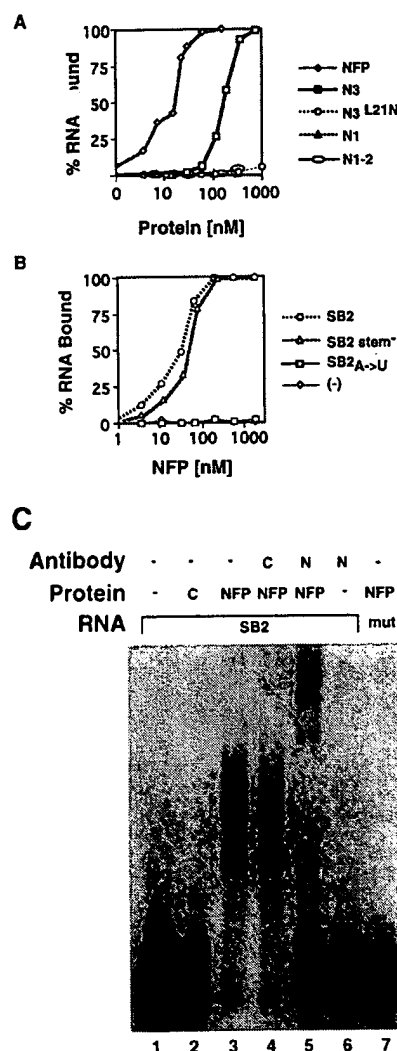


FIG. 2. Nova-1 is a sequence-specific RBP. (A) Nova-1 binding to an RNA selection ligand (SB2). RNA binding of full-length and truncated Nova-1 proteins was determined in 0.5 M LiCl by a filter binding assay (see Materials and Methods). Full-length NFP and the N3 domain proteins bound with  $K_d$ s of ~20 and 100 nM, respectively. Proteins N1, N1-2, and N3 L213N show no binding. (B) Sequence-specific binding of Nova-1 to the RNA selection ligand. Filter binding assays were used to compare the binding of NFP to the wild-type RNA selection ligand SB2 or SB2-derived mutant RNAs. In SB2A3U RNA, the UCAU repeats have been mutated to UC UU, and in SB2 stem<sup>-</sup> RNA, the base pairing of the stem has been disrupted by oligonucleotide-directed mutagenesis (the 3' SB2 stem<sup>-</sup> RNA sequence is UCAU CAAG GGUCUAACGGCGCU, where the third UCAU of SB2 is in boldface and SB2 mismatches are underlined). There is no detectable binding to an irrelevant RNA transcribed from a random library clone (-). (C) Gel shift analysis of Nova-1 protein. SB2 RNA or SB2A3U mutant (mut) <sup>32</sup>P-labeled RNA was transcribed in vitro and incubated with 1 pmol of NFP or irrelevant protein (C), and complexes were run on nondenaturing polyacrylamide gels. These complexes were supershifted with affinity-purified rabbit anti-Nova-1 antibodies (N) but not with preimmune serum (C). No significant amounts of label were evident in the wells of any gel-shifted reaction (data not shown).

RNA ligand is a stem-loop RNA; the loop sequence is the major determinant of Nova-1 binding, while the stem is not essential for binding.

To further define the sequence requirements for high-affinity Nova-1 RNA binding, we generated small RNAs consisting

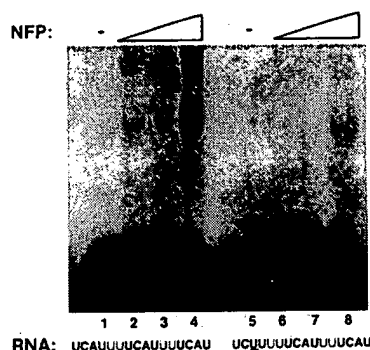


FIG. 3. Gel shift analysis of [UCAU(N)<sub>0-2</sub>]<sub>3</sub> or a mutant loop RNAs bound to NFP. Increasing amounts of NFP (0, 0.5, 1.5, and 5 pmol, respectively, in lanes 1 to 4 and 5 to 8) were incubated with 200 fmol of <sup>32</sup>P-labeled RNA and analyzed by nondenaturing polyacrylamide gel electrophoresis. The RNA used in lanes 1 to 4 consists of the loop sequence indicated below the lanes. The RNA used in lanes 5 to 8 consists of the loop sequence with a single A3U mutation and is indicated below the lanes. UCAU repeats in the RNAs are in boldface. The mutated nucleotide is underlined. All loop RNAs contain a 5' leader sequence (AGG) needed for efficient T7 transcription.

of the [UCAU(N)<sub>0-2</sub>]<sub>3</sub> loop sequence or a series of mutant loop RNAs and assayed their binding to Nova-1 by gel shift and filter binding assays. Gel shift assays with the [UCAU(N)<sub>0-2</sub>]<sub>3</sub> loop RNA revealed a shifted RNA species present at the lowest amount NFP tested (0.5 pmol) that increased with increasing NFP concentration. Gel shift assays with an RNA in which a single UCAU repeat was mutated to UCUU yielded a shifted species only with a 10-fold-higher concentration of NFP (Fig. 3). These data also demonstrate that the NFP-loop sequence RNA (a 21-mer) complex migrated predominantly as a single species, in contrast to migration of NFP-SB2 RNA complex, which migrated as multiple shifted species (Fig. 2C). These data suggest that the isolated loop RNA is unable to support protein multimerization and consequently that Nova-1 multimerization is not essential for high-affinity binding; alternatively, the multiple bands in our gel shift (Fig. 2C) could reflect heterogeneity in RNA secondary structure. Taken together, these results confirm that NFP binds the [UCAU(N)<sub>0-2</sub>]<sub>3</sub> RNA loop sequence specifically and with high affinity.

We next assayed the binding of NFP and N3 to a series of [UCAU(N)<sub>0-2</sub>]<sub>3</sub> loop RNAs harboring specific nucleotide substitutions (Table 2). An A-to-U residue substitution within any one of the three UCAU repeats reduced binding approximately ninefold, consistent with results of the gel shift assay using this mutant (Fig. 3 and data not shown). Substitution of two of three A's with U's completely abolished binding (Table 2). This result demonstrates that Nova-1 is not simply a polypurine tract binding protein but requires specific purine residues to be present within a pyrimidine-rich sequence. Substitution of any of the C, A, or second U nucleotides within each of the three repeats completely eliminated binding, while substitution of the first U within each repeat reduced binding 7- to 12-fold (Table 2). Although the selected RNAs preferentially harbor pyrimidines between UCAU repeats (Fig. 1B), these sequences could be substituted by A residues with no significant change in binding affinity (Table 2). Thus, Nova-1 RNA binding is sequence specific: three intact UCAU repeats are necessary and sufficient for high-affinity binding.

We used the consensus RNA selection sequence to search GenBank by using the BLAST server, to search a database of neuron-specific alternatively spliced exons (59), and to search our own Nova-1 genomic sequence data. These searches ini-

TABLE 2. Nova-1 binding to mutant RNA ligands<sup>a</sup>

RNA	<i>K<sub>d</sub></i> (nM)	
	NFP	N3
<b>UCAU</b> UU <b>UCAU</b> UU <b>UCAU</b> UU	50	230
<b>C</b> --- <b>C</b> --- <b>C</b> --- <b>C</b> ---	340	—
<b>G</b> --- <b>G</b> --- <b>G</b> --- <b>G</b> ---	600	—
<b>-U</b> --- <b>-U</b> --- <b>-U</b> --- <b>-U</b> ---	—	—
<b>-G</b> --- <b>-G</b> --- <b>-G</b> --- <b>-G</b> ---	—	—
<b>--G</b> --- <b>--G</b> --- <b>--G</b> --- <b>--G</b> ---	—	—
<b>---C</b> --- <b>---C</b> --- <b>---C</b> --- <b>---C</b> ---	—	—
<b>---G</b> --- <b>---G</b> --- <b>---G</b> --- <b>---G</b> ---	—	—
<b>--U</b> --- <b>--U</b> --- <b>--U</b> --- <b>--U</b> ---	420	—
<b>--U</b> --- <b>--U</b> --- <b>--U</b> --- <b>--U</b> ---	—	—
<b>--U</b> --- <b>--U</b> --- <b>--U</b> --- <b>--U</b> ---	—	—
<b>---AA</b> --- <b>---AA</b> --- <b>---AA</b> --- <b>---AA</b> ---	50	280

<sup>a</sup>The indicated RNAs were in vitro transcribed from DNA oligonucleotides containing T7 RNA polymerase promoter sequences. All loop RNAs also contain a 5' leader sequence (AGG) needed for efficient T7 transcription. The *K<sub>d</sub>* values are indicated for NFP and N3 binding to RNAs in 0.5 M LiCl; dashes indicate no significant binding (*K<sub>d</sub>* >> 5 μM). *K<sub>d</sub>* were determined by filter binding assays.

tially identified only two potential in vivo Nova-1 RNA binding sites, one within an intron of the GlyR α2 pre-mRNA (35) and one within the Nova-1 pre-mRNA itself (Fig. 4A). The GlyR α2 intron has three distinct UCAU repeats, each separated by two nucleotides (three-fourths are pyrimidines). A fourth UCAU repeat overlaps the first and second in a pattern similar to that seen in RNA selection clone SB31 (Fig. 1B). In the mouse Nova-1 intron, there are two UCAU repeats separated by four pyrimidines, followed by five nucleotides (three of which are pyrimidines) and a CCAU repeat. In both genes, the UCAU sequences are present in potential stem-loop structures and are adjacent to alternatively spliced exons (80 bp upstream of GlyR α2 exon 3A [35] and 8 bp downstream of Nova-1 exon H [reference 9 and Fig. 4A]). In Nova-1, the UCAU element lies within a 35-nucleotide intronic sequence which is 95% identical between human and mouse. In both the GlyR α2 and Nova-1 pre-mRNAs, the sequence CAGU is present one nucleotide upstream of the first UCAU, and a pyrimidine-rich stretch (seven of eight nucleotides in GlyR α2 RNA, six of eight in mouse Nova-1 RNA, and seven of nine in human Nova-1 RNA) is present downstream of the UCAU element.

We cloned these regions of the GlyR α2 and Nova-1 genes and in vitro transcribed them for RNA binding studies. Deletion analysis of the glycine receptor RNA binding the Nova-1 binding site to a 50-bp region containing the UCAU motif (data not shown). Nova-1 bound SB2, GlyR α2, and Nova-1 RNAs in identical manners in the filter binding assay (*K<sub>d</sub>* of 20, 20, and 15 nM, respectively, in 0.5 M LiCl [Fig. 4B]). Mutation of the GlyR α2 RNA (UCAU)<sub>3</sub> repeats to (UAAU)<sub>3</sub> eliminated Nova-1 protein binding as did mutation of the Nova-1 pre-RNA (UCAU)<sub>3</sub> repeats to (UAAU)<sub>3</sub> (Fig. 4B). In addition, Nova-1 protein binding to GlyR α2 RNA could be competed by an excess of SB2 RNA but not by the mutant SB2A3U RNA (Fig. 4C). These results demonstrate that Nova-1 binds to both the GlyR α2 and Nova-1 RNAs in vitro with high affinity via the [UCAU(N)<sub>0-2</sub>]<sub>3</sub> motif.

Previously, we had shown that POMA antibodies target the third KH domain of Nova-1 and can disrupt the low-affinity binding (*K<sub>d</sub>* of ~300 nM) of Nova-1 to polyribomoguanine RNA in vitro (10). Having identified the third KH domain of Nova-1 as both necessary and sufficient for sequence-specific binding, we tested whether POMA antibodies could disrupt the sequence-specific interaction between Nova-1 and the

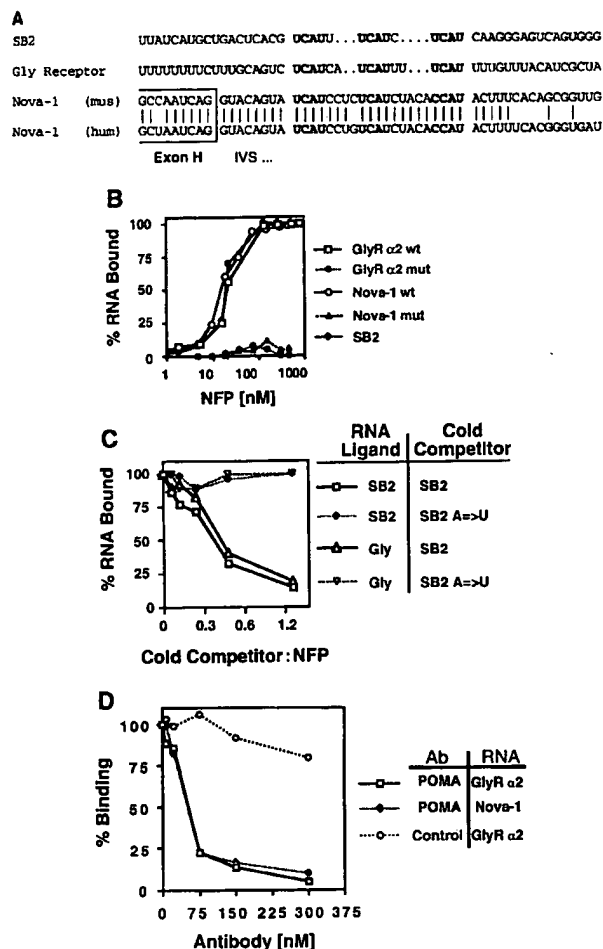


FIG. 4. Nova protein binds to GlyR  $\alpha 2$  and Nova-1 RNAs, and binding can be inhibited by paraneoplastic antibodies. (A) Alignment of *in vivo* [UCAU(N)]<sub>0-23</sub> elements with SB2. UCAU elements present in the rat GlyR  $\alpha 2$  (35) and Nova-1 genomic DNA sequences are shown aligned with UCAU elements (boldface) present in RNA selection clone SB2 (from Fig. 1B), with gaps indicated by dots. Nova-1 genomic sequences were determined from human (hum) and mouse (mus) genomic libraries. The Nova-1 UCAU elements are present within an intron (IVS) 8 nucleotides downstream of an alternatively spliced exon (exon H [9]), and the GlyR  $\alpha 2$  sequences are present ~80 nucleotides upstream of alternatively spliced exon 3A (35). (B) NFP binds to [UCAU(N)]<sub>0-23</sub> elements in GlyR  $\alpha 2$  and Nova-1 RNAs *in vitro*. Fragments of the GlyR  $\alpha 2$  and Nova-1 genes encoding putative Nova-1 binding elements were transcribed *in vitro* and used in filter binding assays with NFP. NFP bound to GlyR  $\alpha 2$  (squares), Nova-1 (open circles), and SB2 RNAs (diamonds) was saturable with a  $K_d$  ~20 nM (standard deviation, ~2 nM). Mutation of the UCAU elements in either GlyR  $\alpha 2$  (closed circles) or Nova-1 RNAs (triangles) eliminates binding. wt, wild type; mut, mutant. (C) NFP binding to the GlyR  $\alpha 2$  can be competed with SB2 RNA. NFP (1.6-pmol) was incubated with 50 fmol of radiolabeled ligand (GlyR  $\alpha 2$  or SB2) and increasing amounts of unlabeled competitor RNA (SB2 or SB2A3U). NFP binding to either SB2 (open squares) or the GlyR  $\alpha 2$  RNA (open triangles) was 50% competed by a ~1:2 molar ratio of unlabeled competitor to NFP, which corresponds to a ~12-fold molar excess of unlabeled SB2 relative to radiolabeled ligand. An excess of cold SB2A3U mutant RNA failed to compete binding of NFP to either radiolabeled SB2 (closed diamonds) or GlyR  $\alpha 2$  (inverted triangle). (D) NFP binding to GlyR  $\alpha 2$  and Nova-1 pre-mRNAs can be inhibited by disease antibodies (Ab). Preincubation of NFP with increasing amounts of affinity-purified disease antibodies (POMA) leads to complete inhibition of RNA binding. Preincubation with similar concentrations of affinity-purified cerebellar degeneration antibodies (Control) had no effect on RNA binding.

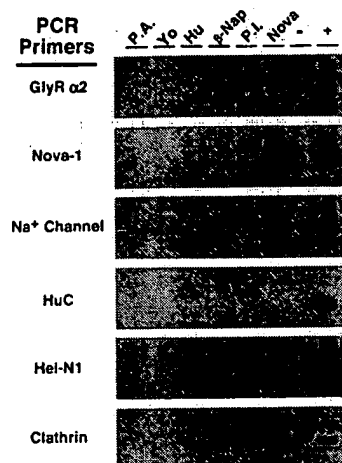


FIG. 5. Nova protein coimmunoprecipitates with GlyR  $\alpha 2$  and Nova-1 pre-mRNAs. Immunoprecipitations of mouse brain were performed as described previously (60) with affinity-purified rabbit anti-Nova-1 antibodies (Nova) or controls, including protein A-Sepharose (P.A.), paraneoplastic cerebellar degeneration serum (Yo), paraneoplastic antiserum directed against the Hu family of n-RBPs (Hu), affinity-purified rabbit anti- $\beta$ -NAP antibodies, or preimmune serum (P.I.) as indicated. RNA present in the precipitates was phenol extracted as described previously (60) and subjected to RT-PCR amplification with the indicated primers. Nova and GlyR  $\alpha 2$  primers bordered the putative Nova-1 binding site, and all PCR products were between 100 and 200 nucleotides in length (see Materials and Methods). RNA present in each immunoprecipitate was also mock amplified without reverse transcriptase (shown only for the Nova immunoprecipitate; -) as a control to ensure the RNA dependence of RT-PCR product. Genomic DNA was amplified (+) to ensure primer fidelity. The fidelity of each immunoprecipitation was confirmed by Western blot analysis (data not shown).

high-affinity GlyR  $\alpha 2$  or Nova-1 RNA ligand. Preincubation of Nova-1 protein with affinity-purified POMA antibodies strongly inhibited Nova-1 binding to both RNAs; a threefold molar excess of antibody reduced RNA binding 80% (Fig. 4D). Nova-1 binding to GlyR  $\alpha 2$  was unaffected by the presence of irrelevant autoimmune antibodies, and incubation of RNA with antibodies alone did not lead to RNA degradation (data not shown). Thus, POMA antibodies are capable of disrupting the high-affinity interaction of Nova-1 and its RNA ligands.

Immunoprecipitation in combination with nucleic acid sequence analysis permits the identification of specific *in vivo* ligands for nucleic acid binding proteins. In previous studies, this method has been used to identify the RNA ligands of autoimmune target antigens and DNA elements bound by transcription factors (see Discussion). To determine whether Nova-1 is bound to GlyR  $\alpha 2$  and Nova-1 pre-mRNAs *in vivo*, we immunoprecipitated crude nuclear extracts of mouse brain with affinity-purified rabbit anti-Nova-1 antibodies. Immunoprecipitates were phenol extracted, and RNA was reverse transcribed and PCR amplified with primers specific for GlyR  $\alpha 2$ , Nova-1, or a series of abundantly expressed neuronal pre-mRNAs. Nova-1-immunoprecipitated RNA gave strong RT-dependent PCR products with both GlyR  $\alpha 2$  and Nova-1 primers but not with control primers (Fig. 5). Control immunoprecipitation with antiserum that recognizes a different family of abundantly expressed n-RBPs (Hu) yielded no detectable bound GlyR  $\alpha 2$  RNA or Nova-1 RNA. Similarly, immunoprecipitations with preimmune rabbit serum, normal human serum, paraneoplastic cerebellar degeneration anti-Yo antibodies, or an irrelevant affinity-purified rabbit antibody to an abundant neuronal autoantigen (anti- $\beta$ -NAP [45]) failed to immunoprecipitate neuronal RNAs. These immunoprecipitations demon-

strate that Nova-1 binds specifically to the GlyR  $\alpha 2$  and Nova-1 pre-mRNAs in vivo and offer an opportunity to identify additional in vivo ligands.

## DISCUSSION

**Nova-1 is a sequence-specific KH-type RBP.** We have shown that Nova-1, a neuron-specific KH-type RBP (9, 10), binds to stem-loop RNAs in a sequence-specific manner in vitro and in vivo. By using affinity elution and stringent binding conditions in our RNA selection protocol, we have identified relatively long sequence-specific pyrimidine-rich RNAs that bind Nova-1 with nanomolar affinity. At least 11 nucleotides in the consensus loop sequence are necessary and sufficient for high-affinity binding, while the stem element confers approximately a three-fold increase in binding affinity but is not essential for binding. Although the high divalent cation concentration present in our RNA selection binding buffer may have promoted selection of RNA ligands harboring stem elements, we note that a number of RBPs bind stem-loop structures both in vitro and in vivo. These include the interactions between Rev and the Rev-responsive element (27, 31) and between the U1-A protein and U1A RNA (51, 52, 64) and the interactions between the L32 RNA (68) and Epstein-Barr virus RNA 1 (17, 63) ligands and their ribosomal proteins (L32 and L22, respectively). In addition, the two in vivo Nova-1 RNA ligands that we have identified are both surrounded by potential stem structures, as assessed by the Zuker RNA folding algorithm ( $\Delta G \approx -15$  [data not shown and reference 71]).

The length of the loop sequence element identified for Nova-1 compares favorably with those of core consensus RNA ligands identified for other RBPs, which typically consist of six to nine nucleotides (12, 29, 36, 62). Based on our mutagenesis of the consensus in vitro Nova-1 RNA ligand and the sequences of the in vivo Nova-1 RNA ligands, at least 11 specific pyrimidine-rich nucleotides present in three repeats appear to be strictly necessary for high-affinity binding. These repeats are interspersed by sequences varying in length that are preferentially (~75%) pyrimidines in both the in vitro RNA selection clones and the in vivo Nova-1 and GlyR  $\alpha 2$  pre-mRNAs, suggesting a Nova-1 RNA binding consensus sequence of [UCAU(Y)<sub>0-2</sub>UCAU(Y)<sub>0-4</sub>NCAU], although a more conservative consensus eliminating the pyrimidine bias is [UCAU(N)<sub>0-4</sub>]<sub>3</sub>. The length of this sequence suggests a limited set of potential in vivo targets, on the order of  $1/4^{11}$  (with no interspersed pyrimidines). Considered in the context of the length of the mammalian genome ( $\sim 3 \times 10^9$  nucleotides), the length of the Nova-1 consensus sequence suggests that there are on the order of tens to perhaps hundreds of possible high-affinity binding sites within the genome. This limited target range increases the relevance of the in vivo Nova-1 RNA targets that we have identified.

KH domains are thought to be involved in RNA binding (10, 11, 56). A large number of KH-type RBPs have been identified since the motif was recognized in 1993 (57). These include the mammalian hnRNP proteins (e.g., hnRNP K) that are thought to be important for the processing of RNA (19) and FMR-1, the fragile-X gene, for which loss of function mutations lead to severe mental retardation (58). However, none have previously been demonstrated to bind RNA with sequence specificity. The functions of a number of KH domain-containing proteins in yeast and invertebrates have been described. MER-1 is a yeast RBP that mediates meiosis-specific alternative splicing of the *MER-2* gene, and PSI is a *Drosophila* protein that inhibits formation of a functional splice variant of the P-element transposase in somatic tissues (54). More recently, the mammalian

splicing factors SF-1 (2) and KSR (41) have been cloned and found to contain one and four KH domains, respectively. In addition, the *Drosophila* Bicaudal-C (Bic-C) and the *Caenorhabditis elegans* MEX-3 KH domain-containing proteins have been suggested to have roles in RNA localization. In several of these instances, the KH domains themselves are thought to play critical roles in the function of these RBPs. A point mutation (I367N) within the second KH domain of FMR-1 results in severe mental retardation, and mutations within the conserved regions of KH domains of *gld-1*, a gene required for oocyte development in *C. elegans*, MEX-3, and Bic-C lead to loss of protein function (18, 32, 39). It is thought that such mutations within the KH domain alter exposed residues within a flexible loop that may directly contact RNA (42, 44). Our observations that the N3 Nova-1 construct containing only the third KH domain of Nova-1 is sufficient to mediate binding to SB2 and that a point mutation (analogous to the FMR-1 mutation) within KH3 abrogates this binding (Fig. 2A) demonstrate that the Nova-1 KH3 domain mediates sequence-specific RNA binding. However, our results do not rule out the possibility that additional residues present in our N3 construct outside the 36- to 38-amino-acid KH3 domain also interact with RNA ligands or that residues absent from our N3 construct (e.g., in the spacer region upstream of KH3) are involved in high-affinity RNA binding (Fig. 2A).

**Identification of in vivo Nova-1 RNA ligands.** We have found evidence that the Nova-1 protein interacts with GlyR  $\alpha 2$  and Nova-1 pre-mRNAs in vivo by coimmunoprecipitation and RT-PCR analysis. In previous studies, protein-RNA coimmunoprecipitation has been used to purify and identify the RNA ligands of numerous autoimmune target antigens (e.g., snRNPs, Ro, and La [60]). This approach has also been used to identify DNA elements bound by transcription factors (e.g., ultrabiothorax, thyroid hormone receptor, and Myc/Max DNA binding elements [7, 23, 24]). Presumably as a result of the low-abundance ligands bound to Nova-1, we were unable to directly identify bound RNA ligands (8a) but were able to analyze candidate RNA ligands by RT-PCR analysis.

The biology of the candidate Nova-1 RNA ligands (GlyR  $\alpha 2$  and Nova-1 pre-mRNAs) are consistent their binding Nova-1 protein in neurons. Both RNAs are expressed in Nova-1-expressing cells. Expression studies demonstrate that GlyR  $\alpha 2$  mRNA is expressed in many brain regions which coexpress Nova-1, including midbrain, brainstem, and spinal cord in embryonic development and continuing into the adult (4). Single-cell studies reveal that Nova-1 protein and GlyR  $\alpha 2$  mRNA are coexpressed within motor neurons in the spinal cord (9, 46, 63a). An interaction between Nova-1 protein and its own pre-mRNA would suggest that Nova-1, like many other RBPs in *Drosophila* and mammals (6, 8, 47), may regulate posttranscriptional processing of its own RNA. Our finding of specific polypyrimidine-rich RNA elements to which Nova-1 binds in vivo suggests potential functions for the Nova-1 protein. While polypyrimidine binding proteins have been implicated in the regulation of polyadenylation, translation, and stability (43), their best-characterized role is in the regulation of pre-mRNA splicing. The ubiquitous polypyrimidine tract binding protein U2AF is an essential splicing factor that binds to polypyrimidine tracts upstream of exon splice acceptors to recruit the U2 snRNP to the splice acceptor site (67). In *Drosophila*, the *skl* protein binds to a specific polypyrimidine-rich RNA that is interspersed with several purine residues (55). Poor polypyrimidine tracts upstream of alternate splice sites in the *tra* pre-mRNA bind relatively weakly to U2AF, which is displaced by *skl*, leading to alternate splice site selection (67). Similarly, polypyrimidine tract binding protein is believed to function as

a splice inhibitor by binding polypyrimidine tracts (37). It is noteworthy that the (UCAUY)<sub>3</sub> elements in both the GlyR  $\alpha$ 2 (35) and Nova-1 (9) pre-mRNAs are adjacent to exons that undergo alternative splicing in neurons and that a number of KH domain-containing RBPs in invertebrates and mammals act as alternative splicing regulators. Taken together, these data suggest that one role of Nova-1 may be to regulate alternative splicing in neurons.

**Nova-1 and neurologic disease.** The interaction of Nova-1 with an inhibitory glycine neurotransmitter receptor RNA may be important in the development of the paraneoplastic neurologic disorder. The neurologic symptoms of POMA are attributable to a loss of motor inhibition (15, 16, 38). Naturally occurring mutations of members of the glycine receptor family in both humans and mice (49, 50) lead to myoclonic neurologic symptoms similar to those seen in POMA. In hereditary hyperekplexia, a human myoclonic neurologic disorder, various point mutations have been found in GlyR  $\alpha$ 1 (53). Similarly, the spastic and spasmodic mice have a myoclonic phenotype and mutations within glycine receptor genes; spasmodic mice have point mutations in GlyR  $\alpha$ 1, and spastic mice have splicing defects due to transposable element insertion within a splice junction of GlyR 2B (49, 50). In addition, expression of a wild-type GlyR 2B transgene in spastic mice rescues the myoclonic phenotype (26). These observations suggest the possibility that the neurologic disease in POMA results from aberrant regulation of glycine receptor expression.

The observation that Nova-1 binding to GlyR  $\alpha$ 2 RNA in vitro is abrogated by POMA disease antibodies (Fig. 4D) suggests a potential mechanism of neuronal dysfunction in these patients: binding of POMA antibody to Nova-1 protein might disrupt the interaction between the Nova-1 protein and GlyR  $\alpha$ 2 pre-mRNA. Since Nova-1 recognizes GlyR  $\alpha$ 2 pre-mRNA sequences upstream from the alternatively spliced exon 3A, disruption of Nova-1 function could result in altered ratios of the mutually exclusive exons 3A and 3B. Although no specific function has been assigned to these alternatively spliced exons, they encode highly related 22-amino-acid domains present on the extracellular region of the receptor, suggesting that these domains may modify receptor-ligand interactions. Such a mechanism presupposes that anti-Nova antibodies are able to gain access to neurons, a process for which there is no direct evidence. However, Nova antibodies are present in high titer in the cerebrospinal fluid of POMA patients (9, 38). Moreover, some data have suggested that autoantibodies can penetrate living cells, including reports that antiribonucleoprotein antibodies may reach intranuclear antigens and that antineuronal antibodies can penetrate neurons (reviewed in references 1 and 16).

#### ACKNOWLEDGMENTS

We thank Nat Heintz, Titia de Lange, and Magda Konarska for helpful and stimulating discussions and Nat Heintz, Titia de Lange, and members of our laboratory for critical reading of the manuscript. We also thank Yolanda Yang for assistance developing the filter binding assays.

This work was supported by the NINDS (grant RO1 NS34389) and a Hirschl Career Scientist Award to R.B.D. R.J.B. was supported by NIH MSTP training grant 5T32 GM07739.

#### REFERENCES

- Alarcon-Segovia, D., A. Ruiz-Argüelles, and L. Llorente. 1996. Broken dogma: penetration of autoantibodies into living cells. *Immunol. Today* 17:163-164.
- Arning, S., P. Grüter, G. Bilbe, and A. Kramer. 1996. Mammalian splicing factor SF1 is encoded by variant cDNAs and binds to RNA. *RNA* 2:794-810.
- Bartel, D. P., M. L. Zapp, M. R. Green, and J. W. Szostak. 1991. HIV-1 Rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* 67:529-536.
- Bechade, C., C. Sur, and A. Triller. 1994. The inhibitory neuronal glycine receptor. *Bioessays* 16:735-744.
- Bell, L., E. Maine, P. Schedl, and T. Cline. 1988. Sex-lethal, a *Drosophila* sex determination switch gene, exhibits sex-specific RNA splicing and sequence similarity to RNA binding proteins. *Cell* 55:1037-1046.
- Bell, L. R., J. I. Horabin, P. Schedl, and T. W. Cline. 1991. Positive autoregulation of sex-lethal by alternative splicing maintains the female determined state in *Drosophila*. *Cell* 65:229-240.
- Bigler, J., and R. N. Eisenman. 1994. Isolation of a thyroid hormone-responsive gene by immunoprecipitation of thyroid hormone receptor-DNA complexes. *Mol. Cell. Biol.* 14:7621-7632.
- Boelens, W., E. Jansen, W. van Venrooij, R. Stripecte, I. Mattaj, and S. Gunderson. 1993. The human U1 snRNP-specific U1A protein inhibits polyadenylation of its own pre-mRNA. *Cell* 72:881-892.
- Buckanovich, R. J., and R. B. Darnell. Unpublished data.
- Buckanovich, R. J., J. B. Posner, and R. B. Darnell. 1993. Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding protein and is specifically expressed in the developing motor system. *Neuron* 11:657-672.
- Buckanovich, R. J., Y. Y. Yang, and R. B. Darnell. 1996. The onconeural antigen Nova-1 is a neuron-specific RNA binding protein, the activity of which is inhibited by paraneoplastic antibodies. *J. Neurosci.* 16:1114-1122.
- Burd, C. G., and G. Dreyfuss. 1994. Conserved structures and diversity of functions of RNA-binding proteins. *Science* 265:615-621.
- Burd, C. G., and G. Dreyfuss. 1994. RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. *EMBO J.* 13:1197-1204.
- Caceres, J. F., S. Stamm, D. M. Helfan, and A. R. Krainer. 1994. Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. *Science* 265:1706-1709.
- Cullen, B. R. 1994. RNA-sequence-mediated gene regulation in HIV-1. *Infect. Agents Dis.* 3:68-76.
- Darnell, R. B. 1994. Paraneoplastic syndromes, p. 137-141. *In* E. Feldmann (ed.), *Current diagnosis in neurology*. Mosby-Year Book, Inc., Philadelphia, Pa.
- Darnell, R. B. 1996. Onconeural antigens and the paraneoplastic neurologic disorders: at the intersection of cancer, immunity and the brain. *Proc. Natl. Acad. Sci. USA* 93:4529-4536.
- Darnell, R. B. Unpublished data.
- Dobbelstein, M., and T. Shenk. 1995. In vitro selection of RNA ligands for the ribosomal L22 protein associated with Epstein-Barr virus-expressed RNA by using randomized and cDNA-derived RNA libraries. *J. Virol.* 69:8027-8034.
- Draper, B. W., C. C. Mello, B. Bowerman, J. Hardin, and J. R. Priess. 1996. MEX-3 is a KH domain protein that regulates blastomere identity in early *C. elegans* embryos. *Cell* 87:205-216.
- Dreyfuss, G., M. J. Matunis, S. Piñol-Roma, and C. G. Burd. 1993. hnRNP proteins and the biogenesis of mRNA. *Annu. Rev. Biochem.* 62:289-321.
- Ellington, A., and J. Szostak. 1990. In vitro selection of RNA molecules that bind specific ligands. *Nature* 346:818-822.
- Engbrecht, J. A., K. Voelkel-Meiman, and G. S. Roeder. 1991. Meiosis-specific RNA splicing in yeast. *Cell* 66:1257-1268.
- Famulok, M., and J. W. Szostak. 1992. Stereospecific recognition of tryptophan agarose by in vitro selected RNA. *J. Am. Chem. Soc.* 114:3990-3991.
- Gould, A. P., J. J. Brookman, D. I. Strutt, and W. R.A.H. 1990. Targets of homeotic gene control in *Drosophila*. *Nature* 348:308-311.
- Grandori, C., J. Mac, F. Siebelt, D. E. Ayer, and R. N. Eisenman. 1996. Myc-max heterodimers activate a DEAD box gene and interact with multiple E box-related sites *in vivo*. *EMBO J.* 15:4344-4357.
- Green, R., A. D. Ellington, D. P. Bartel, and J. W. Szostak. 1991. *In vitro* genetic analysis: selection and amplification of rare functional nucleic acids. *Methods (Orlando)* 2:75-86.
- Hartenstein, B., J. Schenkel, J. Kuhse, B. Besenbeck, C. Kling, C. M. Becker, H. Betz, and H. Weiher. 1996. Low level expression of glycine receptor beta subunit transgene is sufficient for phenotype correction in spastic mice. *EMBO J.* 15:1275-1282.
- Heaphy, S., C. Dingwall, I. Ernberg, M. J. Gait, S. Green, J. Karn, A. D. Lowe, M. Singh, and M. A. Skinner. 1990. HIV-1 regulator of virion expression (rev) protein binds to an RNA stem-loop structure located within the rev responsive element region. *Cell* 60:685-693.
- Hedley, M. L., and T. Maniatis. 1991. Sex-specific splicing and polyadenylation of *dsx* pre-mRNA requires a sequence that binds specifically to *tra-2* protein *in vitro*. *Cell* 65:579-586.
- Heinrichs, V., and B. S. Baker. 1995. The *Drosophila* SR protein RBP1 contributes to the regulation of doublesex alternative splicing by recognizing RBP1 RNA target sequences. *EMBO J.* 14:3987-4000.
- Inoue, K., K. Hoshijima, H. Sakamoto, and Y. Shimura. 1990. Binding of the *Drosophila* sex-lethal gene product to the alternative splice site of trans-former primary transcript. *Nature* 344:461-463.
- Jensen, K. B., L. Green, S. MacDougall-Waugh, and C. Tuerk. 1994. Char-

- acterization of an in vitro-selected RNA ligand to the HIV-1 Rev protein. *J. Mol. Biol.* 235:237-247.
32. Jones, A. R., and T. Schedl. 1995. Mutations in *glid-1*, a female germ cell-specific tumor suppressor gene in *Caenorhabditis elegans*, affect a conserved domain also found in Src-associated protein Sam68. *Genes Dev.* 9:1491-1504.
  33. Kamma, H., D. S. Portman, and G. Dreyfuss. 1995. Cell type-specific expression of hnRNP proteins. *Exp. Cell Res.* 221:187-196.
  34. Kjems, J., M. Brown, D. D. Chang, and P. A. Sharp. 1991. Structural analysis of the interaction between the human immunodeficiency virus Rev protein and the Rev response element. *Proc. Natl. Acad. Sci. USA* 88:683-687.
  35. Kuhse, J., A. Kuryatov, Y. Maulet, M. L. Malosio, V. Schmieden, and H. Betz. 1991. Alternative splicing generates two isoforms of the  $\alpha 2$  subunit of the inhibitory glycine receptor. *FEBS Lett.* 283:73-77.
  36. Levine, T. D., F. Gao, P. H. King, L. G. Andrews, and J. D. Keene. 1993. Hel-N1: an autoimmune RNA-binding protein with specificity for 3' uridylic-rich untranslated regions of growth factor mRNAs. *Mol. Cell. Biol.* 13:3494-3504.
  37. Lin, C.-H., and J. Patton. 1995. Regulation of alternative 3' splice site selection by constitutive splicing factors. *RNA* 1:234-245.
  38. Luque, F., H. Furneaux, R. Ferziger, M. Rosenblum, S. Wray, S. Schold, M. Glantz, K. Jaekle, H. Biran, M. Lesser, W. Paulsen, M. River, and J. Posner. 1991. Anti-Ri: an antibody associated with paraneoplastic opsoclonus and breast cancer. *Ann. Neurol.* 29:241-251.
  39. Mahone, M., E. E. Saffman, and P. F. Lasko. 1995. Localized *Bicaudal-C* RNA encodes a protein containing a KH domain, the RNA binding motif of FMR1. *EMBO J.* 9:2043-2055.
  40. Malim, M., J. Hauber, S. Y. Le, J. Maizel, and B. Cullen. 1989. The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* 338:254-257.
  41. Min, H., C. W. Turck, and D. L. Black. A new regulatory protein, KSR, mediates exon inclusion through an intronic splicing enhancer. Submitted for publication.
  42. Morelli, M. A. C., G. Stier, T. Gibson, C. Joseph, G. Musco, A. Pastore, and G. Trave. 1995. The KH module has an  $\alpha\beta$  fold. *FEBS Lett.* 358:193-198.
  43. Morris, D. R., T. Kakegawa, R. L. Kaspar, and M. W. White. 1993. Polypyrimidine tracts and their binding proteins: regulatory sites for post-transcriptional modulation of gene expression. *Biochemistry* 32:2931-2937.
  44. Musco, G., G. Stier, C. Joseph, M. A. C. Morelli, M. Nilges, T. Gibson, and A. Pastore. 1996. Three-dimensional structure and stability of the KH domain: molecular insights into the fragile X syndrome. *Cell* 85:237-245.
  45. Newman, L. S., M. O. McKeever, H. J. Okano, and R. B. Darnell. 1995.  $\beta$ -NAP, a cerebellar degeneration antigen, is a neuron-specific vesicle coat protein. *Cell* 82:773-783.
  46. Racca, C., A. Gardiol, and A. Triller. 1997. Dendritic and postsynaptic localizations of glycine receptor  $\alpha$  subunit mRNAs. *J. Neurosci.* 17:1691-1700.
  47. Rongo, C., E. R. Gavis, and R. Lehmann. 1995. Localization of oskar RNA regulates oskar translation and requires oskar protein. *Development* 121:2737-2746.
  48. Ruskin, B., P. D. Zamore, and M. R. Green. 1988. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell* 52:207-219.
  49. Ryan, S. G., M. S. Buckwalter, J. W. Lynch, C. A. Handford, L. Segura, R. Shiang, J. J. Wasmuth, S. A. Camper, P. Schofield, and P. O'Connell. 1994. A missense mutation in the gene encoding the  $\alpha 1$  subunit of the inhibitory glycine receptor in the spasmodic mouse. *Nat. Genet.* 7:131-135.
  50. Saul, B., V. Schmieden, C. Kling, C. Mulhardt, P. Gass, J. Kuhse, and C. M. Becker. 1994. Point mutation of glycine receptor  $\alpha 1$  subunit in the spasmodic mouse affects agonist responses. *FEBS Lett.* 350:71-76.
  51. Scherly, D., W. Boelens, N. A. Dathan, W. J. van Venrooij, and I. W. Mattaj. 1990. Major determinants of the specificity of interaction between small nuclear ribonucleoproteins U1A and U2B and their cognate RNAs. *Nature* 345:502-506.
  52. Scherly, D., W. Boelens, W. J. van Venrooij, N. A. Dathan, J. Hamm, and I. W. Mattaj. 1989. Identification of the RNA binding segment of human U1 A protein and definition of its binding site on U1 snRNA. *EMBO J.* 8:4163-4170.
  53. Shiang, R., S. G. Ryan, Y. Z. Zhu, A. F. Hahn, P. O'Connell, and J. J. Wasmuth. 1993. Mutations in the  $\alpha 1$  subunit of the inhibitory glycine receptor cause the dominant neurologic disorder, hyperekplexia. *Nat. Genet.* 5:351-358.
  54. Siebel, C. W., A. Admon, and D. C. Rio. 1995. Soma-specific expression and cloning of PSI, a negative regulator of P element pre-mRNA splicing. *Genes Dev.* 9:269-283.
  55. Singh, R., J. Valcarcel, and M. R. Green. 1995. Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science* 268:1173-1176.
  56. Siomi, H., M. Choi, M. Siomi, R. Nussbaum, and G. Dreyfuss. 1994. Essential role for KH domains in RNA binding: impaired RNA binding by a mutation in the KH domain of FMR1 that causes fragile X syndrome. *Cell* 77:33-39.
  57. Siomi, H., M. J. Matunis, W. M. Michael, and G. Dreyfuss. 1993. The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. *Nucleic Acids Res.* 21:1193-1198.
  58. Siomi, H., M. Siomi, R. Nussbaum, and G. Dreyfuss. 1993. The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein. *Cell* 74:291-298.
  59. Stamm, S., M. Q. Zhang, T. G. Marr, and D. M. Helfman. 1994. A sequence compilation and comparison of exons that are alternatively spliced in neurons. *Nucleic Acids Res.* 9:1515-1526.
  60. Steitz, J. 1989. Immunoprecipitation of ribonucleoproteins using autoantibodies. *Methods Enzymol.* 180:468-481.
  61. Szostak, J. W., and A. D. Ellington. 1993. In vitro selection of functional RNA sequences. The RNA world. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
  62. Tacke, R., and J. L. Manley. 1995. The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities. *EMBO J.* 14:3540-3551.
  63. Toczyski, D. P., and J. A. Steitz. 1993. The cellular RNA-binding protein EAP recognizes a conserved stem-loop in the Epstein-Barr virus small RNA EBER 1. *Mol. Cell. Biol.* 13:703-710.
  - 63a. Triller, A., and R. B. Darnell. Unpublished data.
  64. Tsai, D. E., D. S. Harper, and J. D. Keene. 1991. U1-snRNP-A protein selects a ten nucleotide consensus sequence from a degenerate RNA pool presented in various structural contexts. *Nucleic Acid Res.* 19:4931-4936.
  65. Tuerk, C., and L. Gold. 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249:505-510.
  66. Tuerk, C., and S. MacDougall-Waugh. 1993. In vitro evolution of functional nucleic acids: high-affinity RNA ligands of HIV-1 proteins. *Gene* 137:33-39.
  67. Valcarcel, J., R. Singh, P. Zamore, and M. Green. 1993. The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. *Nature* 362:171-175.
  68. Vilardell, J., and J. R. Warner. 1994. Regulation of splicing at an intermediate step in the formation of the spliceosome. *Genes Dev.* 8:211-220.
  69. Zamore, P. D., J. G. Patton, and M. R. Green. 1992. Cloning and domain structure of the mammalian splicing factor U2AF. *Nature* 355:609-614.
  70. Zapp, M., and M. Green. 1989. Sequence-specific RNA binding by the HIV-1 rev protein. *Nature* 342:714-716.
  71. Zuker, M. 1989. On finding all suboptimal foldings of a RNA molecule. *Science* 244:48-52.

# The Determinants of RNA-binding Specificity of the Heterogeneous Nuclear Ribonucleoprotein C Proteins\*

(Received for publication, May 31, 1994)

Matthias Görlach†, Christopher G. Burd§, and Gideon Dreyfuss¶

From the Howard Hughes Medical Institute and Department of Biochemistry and Biophysics, University of Pennsylvania, School of Medicine, Philadelphia, Pennsylvania 19104-6148

The hnRNP C proteins (C1/C2) are tenacious nuclear pre-mRNA-binding proteins that belong to the large RNP motif family of RNA-binding proteins. This motif identifies an RNA-binding domain (RBD) that consists of a four-stranded antiparallel  $\beta$ -sheet packed against two  $\alpha$ -helices. Despite considerable information on the structure of the hnRNP C RBD, little is known about its RNA-binding properties. To address this we used *in vitro* selection/amplification from pools of random sequence RNA to determine the RNA-binding specificity of hnRNP C1. After 8 rounds of selection/amplification nearly all RNAs contained contiguous stretches of at least 5 U residues, and filter-binding assays demonstrated that this sequence constitutes a high-affinity ( $K_d = 170$  nM) binding site for hnRNP C1. The highest affinity we measured for hnRNP C1 was for r(U)<sub>14</sub> ( $K_d = 14$  nM). An RBD-containing peptide fragment of hnRNP C1 (amino acids 2-94) bound oligoribonucleotides containing an hnRNP C1 high-affinity binding site with nearly equal affinity to that of hnRNP C1. Unlike hnRNP C1, however, this peptide also bound oligoribonucleotides that do not contain high-affinity hnRNP C1-binding sites. We identified a region of 10 amino acids, immediately COOH-terminal to the RNP motif (amino acids 95-104), that prevents the minimal RBD from binding nonspecific RNA ligands. We propose that the highly conserved  $\beta\alpha\beta\alpha\beta$  core structure of the RNP motif RBD confers a general RNA binding activity to RNP motif RBDs and that the determinants of RNA-binding specificity reside in the most variable regions, the loops connecting the  $\beta$ -strands and/or the contiguous NH<sub>2</sub> and COOH termini of the RBD.

Nascent pre-messenger RNA (pre-mRNA), or heterogeneous nuclear (hn)RNA,<sup>1</sup> associates with nuclear RNA-binding proteins upon its emergence from the RNA polymerase II complex. These proteins, which bind pre-mRNA and are not stable components of other nuclear structures (e.g. small nuclear ribonucleoproteins), are collectively termed hnRNP proteins (Drey-

fuss, 1986; Dreyfuss *et al.*, 1993; Görlach *et al.*, 1993). Most, if not all hnRNP proteins bind RNA directly (Cobianchi *et al.*, 1988; Merrill *et al.*, 1988; Piñol-Roma *et al.*, 1988; Swanson and Dreyfuss, 1988a, 1988b; Wilusz *et al.*, 1988; Wilusz and Shenk 1990; Matunis *et al.* 1992, 1993; Bennet *et al.*, 1993) and in doing so, they can influence RNA-processing reactions and, ultimately, gene expression.

The hnRNP C proteins (C1 and C2) are very abundant nuclear proteins, and they are among the most avid pre-mRNA-binding proteins in human (HeLa) cells and several lines of evidence suggest that they bind RNA with sequence specificity (Piñol-Roma *et al.*, 1988; Swanson and Dreyfuss 1988a, 1988b; Wilusz and Shenk, 1990). They bind poly(U) tenaciously and can bind oligo(U) stretches found in natural RNAs (Swanson and Dreyfuss 1988a, 1988b; Wilusz *et al.*, 1988; Wilusz and Shenk 1990). A recent NMR study with the hnRNP C RBD (amino acids 2-94) found that it binds oligo-r(U)<sub>8</sub> with a  $K_d$  in the range of 1-15  $\mu$ M (Görlach *et al.*, 1992).

Many hnRNP proteins, including hnRNP C1 and C2, belong to a family of RNA-binding proteins which share an evolutionary conserved region of 90-100 amino acids termed the RNP motif, the hallmarks of which are the RNP 1 octamer and the RNP 2 hexamer consensus sequences (Dreyfuss *et al.*, 1993) (also called RRM (Query *et al.*, 1989) or RNP 80 (Mattaj, 1989; Bandziulis *et al.*, 1989; Kenan *et al.*, 1991)). The RNP motif RBD has been shown to be a bona fide RNA-binding domain (Query *et al.*, 1989; Scherly *et al.*, 1989, 1990a; Lutz-Freyermuth *et al.*, 1990; Nietfeld *et al.*, 1990; Burd *et al.*, 1991). Recent structural analyses of the NH<sub>2</sub>-terminal RBD of snRNP U1A protein (U1A) and the hnRNP C proteins RBD have shown that the domain has a  $\beta\alpha\beta\alpha\beta$  structure and that it folds into a four-stranded antiparallel  $\beta$ -sheet which is packed against two  $\alpha$ -helices. Amino acids of RNP 1 and RNP 2, which interact with RNA (Merrill *et al.*, 1988; Schwemmler *et al.*, 1989), are juxtaposed on the two central  $\beta$ -strands of the folded RBD (Nagai *et al.*, 1990; Hoffman *et al.*, 1991; Wittekind *et al.*, 1992). Many residues located on the  $\beta$ -sheet (Scherly *et al.*, 1989; Surowy *et al.*, 1989; Nagai *et al.*, 1990; Jessen *et al.*, 1991) and in the flanking amino- and carboxyl-terminal regions probably interact with RNA (Görlach *et al.*, 1992).

An important question concerning RNP motif proteins centers around the elements which confer RNA-binding specificity. A major determinant of U1A and snRNP U2B' (U2B'') RNA-binding specificity is the loop connecting  $\beta$ 2 and  $\beta$ 3 (loop 3). Replacing this loop in U1A with the analogous one of U2B' confers U2B' specificity to U1A (Scherly *et al.*, 1990a). However, U2B' and the hybrid U1A protein need an auxiliary protein (snRNP U2A') to specifically bind the cognate RNA ligand (Scherly *et al.*, 1990a, 1990b). In contrast, the RBD of the hnRNP C1/C2 proteins lacks loop 3 (Swanson *et al.*, 1987; Bandziulis *et al.*, 1989; Burd *et al.*, 1989; Kenan *et al.*, 1991; Wittekind *et al.*, 1992) indicating that it is not essential for the RNA binding activity of all RNP motif proteins.

\* This work was supported by the Howard Hughes Medical Institute, grants from the National Institutes of Health, and in part by a postdoctoral fellowship (to M. G.) by the Deutsche Forschungsgemeinschaft. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Present address: Institut für Molekulare Biotechnologie, 07708 Jena, Germany.

§ Present address: Howard Hughes Medical Institute, University of California, San Diego, La Jolla, CA 92093.

¶ To whom correspondence should be addressed.

<sup>1</sup> The abbreviations used are: hnRNP, heterogeneous nuclear ribonucleoprotein; sn, small nuclear; RBD, RNA-binding domain; ss, single stranded.

In order to better understand the RNA binding activity of hnRNP C1, and RNP motif proteins in general, we determined the RNA-binding specificity of hnRNP C1 by *in vitro* selection/amplification (Tuerk and Gold, 1990). The affinities for different RNA sequences of hnRNP C1 and two protein fragments containing its RBD were determined. The results demonstrate that 10 amino acids immediately carboxyl-terminal to the minimal RBD of hnRNP C1 are primarily responsible for sequence specific RNA binding.

#### MATERIALS AND METHODS

**In Vitro Selection of RNA**—*In vitro* selection/amplification from pools of RNA oligomers randomized in 20 positions (Tuerk and Gold, 1990) with purified hnRNP C1, which was immobilized via the monoclonal antibody 4F4 (Choi and Dreyfuss, 1984) on protein A-Sepharose (Pharmacia Biotech Inc.), were performed as described elsewhere (Burd and Dreyfuss, 1994). Eight rounds of selection/amplification were carried with 30 pmol of hnRNP C1 and a RNA to protein ratio of 100:1 in the first round; this ratio varied in subsequent rounds between 100:1 and 25:1.

**Binding Assays**—Binding of *in vitro* transcribed and translated hnRNP C1 and the RBD polypeptides K94 and M104 to Sepharose-linked poly(rU) was carried out at 1 M KCl as described in Swanson and Dreyfuss (1988a). Filter-binding assays using 5'-<sup>32</sup>P-labeled RNA oligonucleotides were carried out in 20  $\mu$ l of binding buffer (10 mM Hepes-KOH, pH 7.4, 100 mM KCl, and 2.5 mM MgCl<sub>2</sub>) at 30 °C for 30 or 5 min (binding was saturated after 1 min, data not shown). 15  $\mu$ l of the reaction was spotted onto Millipore HAWP 02500 nitrocellulose filters pre-washed with 5 ml of binding buffer. The filters were washed with 1 ml of ice-cold binding buffer once at a negative pressure of 15–20 Hg/square inch. Filter-bound RNA was measured by liquid scintillation counting and the amount of bound RNA was plotted as fraction of the RNA bound at saturation (typically 60–100% of input RNA) against protein concentration. Binding curves derived from at least three independent binding experiments were fitted using the curve fitting routine provided with SigmaPlot (Jandel Scientific, San Rafael, CA). The concentration of protein active in these assays was determined at saturating protein concentration titrating in an excess of RNA (Witherell and Uhlenbeck, 1989). The fitted binding curves which were corrected for the active fraction of protein in the binding assay were used to determine the apparent dissociation constant ( $K_d$ ) which is the concentration of protein required for 50% saturation of binding (Kelly *et al.*, 1976).

**Gel Electrophoresis**—Labeled RNA oligonucleotides were analyzed on 18% polyacrylamide, 7 M urea gels (Sambrook *et al.*, 1989) and NaDodecylSO<sub>4</sub>-polyacrylamide electrophoresis and fluorography was carried out as described (Dreyfuss *et al.*, 1984).

**Cells, DNA, RNA, Protein Overexpression**—Deletional mutants of the hnRNP C1 cDNA and constructs for the overexpression of hnRNP C1 and its RNA-binding domain (RBD) derivatives were derived from pHC12 (Swanson *et al.*, 1987) by using appropriate restriction enzymes or polymerase chain reaction. Heat-inducible overexpression of hnRNP C1 in *Escherichia coli* was achieved using the pRC 23 (Crowl *et al.*, 1985) system. Overexpression of the RBD construct K94 encompassing the first 94 amino acids of hnRNP C1 and expression of M104 which encompasses the first 104 amino acids of hnRNP C1 was carried out in *E. coli* BL21(DE3) with the pET system (Studier *et al.*, 1990) using pET3d for K94 and pET11d for M104. RNA oligonucleotides were synthesized and purified as described (Görlach *et al.*, 1992).

**Protein Purification**—Overexpressed hnRNP C1 was purified from *E. coli* S30 extracts (extraction buffer: 50 mM Hepes-KOH, pH 7.5, 500 mM NaCl, 10 mM Na<sub>2</sub>S<sub>2</sub>O<sub>8</sub>, 1 mM dithiothreitol, 1 mM NaEDTA, 1  $\mu$ g/ml each of leupeptin and pepstatin; 0.5% aprotinin, and 10% (v/v) glycerol). The extract was centrifuged at 30,000  $\times$  g for 30 min and the supernatant was passed through a DEAE-cellulose column and the flow through was directly loaded onto a ssDNA-cellulose column. The ssDNA-cellulose was washed extensively with ssDNA buffer containing 50 mM Na<sub>2</sub>HPO<sub>4</sub>, NaH<sub>2</sub>PO<sub>4</sub>, pH 7.5, 0.5 mM dithiothreitol, 0.2 mM NaEDTA, 10% glycerol, and 500 mM NaCl and pre-eluted with the same buffer containing 1 mg/ml heparin (Piñol-Roma *et al.*, 1988). The hnRNP C1 protein was eluted using ssDNA buffer at 2 M NaCl, concentrated by ultrafiltration to 3 mg/ml protein, and stored at -80 °C; the protein was approximately 95% pure as estimated from Coomassie-stained gels. The K94 RBD and the M104 RBD were purified from *E. coli* S100 extracts using extraction buffer (see above). The extracts were diluted to a NaCl concentration of 100 mM (K94) or 150 mM (M104), respectively, with H<sub>2</sub>O-buffer (10 mM Hepes-KOH, pH 7.5, 0.2 mM NaEDTA, 10% (v/v) glycerol), centrifuged

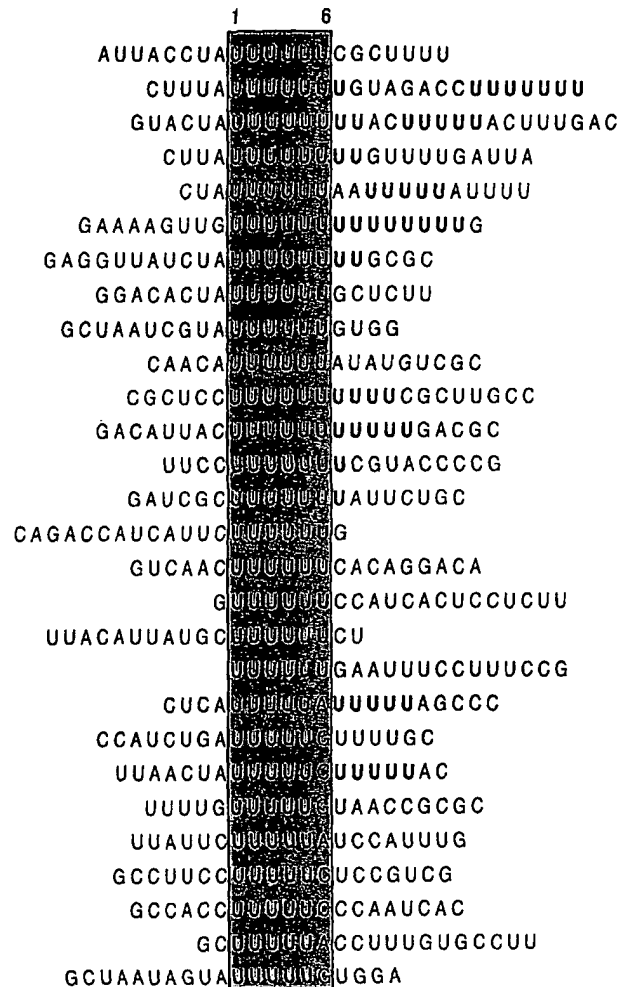
at 100,000  $\times$  g for 1 h and the supernatant was loaded onto a 500-ml DEAE Sephacel column. The flow through of the DEAE-Sepharose column was directly fed into a 90-ml ssDNA-cellulose column equilibrated with H100 (H<sub>2</sub>O buffer with 100 mM NaCl for K94) or H150 (H<sub>2</sub>O buffer with 150 mM NaCl for M104), respectively. The ssDNA-cellulose was extensively washed and then pre-eluted with 1 mg/ml heparin in H100 (K94) or H150 (M104), respectively. K94 and M104 were eluted from ssDNA-cellulose with H500 and dialyzed against H100 buffer. The dialyzed protein samples were loaded onto a 90-ml S-Sepharose column which was developed with a salt gradient from 100 to 600 mM NaCl in H<sub>2</sub>O-buffer. Eluted protein was concentrated by (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> precipitation at 80% saturation dialyzed against H100 without glycerol and further concentrated by ultrafiltration up to a concentration of 30 mg/ml and stored at 4 °C. The K94 and M104 preparations appear to be homogeneous by one- and two-dimensional gel analysis and other criteria (see Wittekind *et al.* (1992)).

#### RESULTS AND DISCUSSION

**RNA-binding Specificity of hnRNP C1**—In order to determine the RNA-binding specificity of hnRNP C1 we performed *in vitro* selection/amplification experiments using pools of RNA oligonucleotides randomized at 20 positions (Tuerk and Gold, 1990). In these experiments recombinant hnRNP C1 was bound to protein A-Sepharose with a monoclonal antibody (4F4; Choi and Dreyfuss (1984)) that reacts with an epitope near the COOH terminus of hnRNP C1.<sup>2</sup> Immobilized hnRNP C1 was then incubated with a large molar excess of random sequence RNA, and bound RNAs were subsequently purified and amplified by reverse transcriptase-polymerase chain reaction using primers to constant sequences outside the randomized region. The resultant DNA was transcribed with T7 RNA polymerase and this process was repeated for a total of 8 cycles (Tuerk and Gold, 1990; Burd and Dreyfuss, 1994). After the final cycle the products were cloned and the sequence of the selected region of 28 clones was determined (Fig. 1). Most of the selected RNAs contain oligo(rU) stretches ranging in length from 5 to 11 U residues, the most prevalent being runs of 5 or 6 uridylate residues (Fig. 1). A second experiment using RNA oligonucleotides randomized at 10 positions yielded essentially the same results (data not shown). The sequences flanking the oligo(rU) stretches exhibit variability and the oligo(rU) stretches do not reside in a defined position with respect to the termini of the 20-mer region, suggesting that the greatest selective constraint is on the oligo(rU) stretch itself (Fig. 1). These experiments suggest that the sequence, UUUUU, is a high-affinity binding site for hnRNP C1.

To confirm that the *in vitro* selection/amplification experiments identified RNA molecules containing a high-affinity hnRNP C1-binding site(s), we measured the apparent equilibrium dissociation constants ( $K_d$ ) of recombinant hnRNP C1 for oligoribonucleotides containing sequences identified in the selection/amplification experiments or unrelated (*i.e.* not selected) sequences. An oligoribonucleotide (14-mer) derived from a selected RNA (CB2, see Fig. 3) which contains a central r(U)<sub>5</sub> stretch was bound with an apparent dissociation constant ( $K_d$ ) of 170 nM (Fig. 3 and Table I). As many of the selected RNAs contain uridine stretches longer than r(U)<sub>5</sub>, we also measured the affinity of hnRNP C1 for r(U)<sub>14</sub>; it was bound with 10-fold higher affinity ( $K_d \sim 14$  nM; Fig. 3 and Table I). To determine the affinity of hnRNP C1 for nonselected RNA sequences, a 20-mer derived from the first intron of the human  $\beta$ -globin gene and r(C)<sub>14</sub> were used. Each of these oligoribonucleotides do not contain sequences that resemble a putative high-affinity hnRNP C1-binding site and they were not significantly bound by hnRNP C1 even at the highest protein concentration tested (Fig. 3, C1, and Table I). These results, as well as earlier UV cross-linking experiments (Wilusz and Shenk, 1990), confirm

<sup>2</sup> C. G. Burd and G. Dreyfuss, unpublished observations.



## CONSENSUS UUUUU

FIG. 1. High affinity substrates of hnRNP C1. Sequences selected by hnRNP C1 from a pool of RNA oligonucleotides containing 20 randomized positions (see "Materials and Methods") are shown. The minimal length stretches of 5 or 6 uninterrupted uridylyte residues are aligned and boxed. The sequence from which the CB2 oligonucleotide used in the binding experiments was derived is the last sequence in this list.

TABLE I  
Apparent dissociation constants ( $K_d$ ) for RNA of hnRNP C1 and its RBD derivatives M104 and K94

RNA ligand	hnRNP C1	M104	K94
CB2	170	470	450
r(U) <sub>14</sub>	14	400	530
$\beta$ -Globin intron	NA <sup>a</sup>	NA	660
r(C) <sub>14</sub>	ND <sup>b</sup>	NA <sup>c</sup>	2100 <sup>c</sup>

<sup>a</sup> NA, not applicable; no binding detected.

<sup>b</sup> ND, not determined; does not bind poly(rC) (Swanson and Dreyfuss, 1988a).

<sup>c</sup> Data not shown.

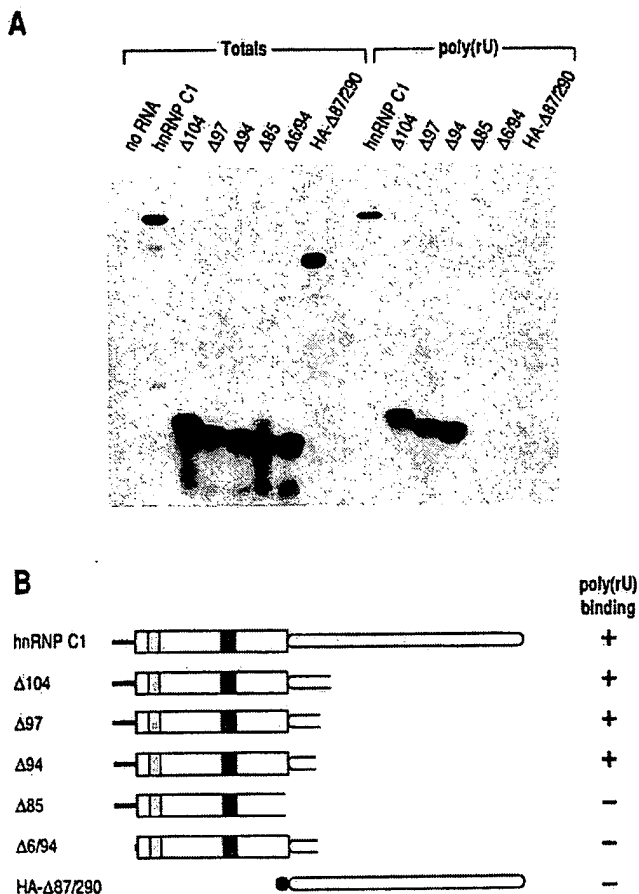
that the sequence, UUUUU, is an hnRNP C1 high-affinity binding site. Importantly, they also demonstrate that hnRNP C1 is a sequence-specific RNA-binding protein. Recently, another abundant hnRNP protein, hnRNP A1, was shown to bind RNA with sequence specificity (Burd and Dreyfuss, 1994) so it seems likely that this is a general property of hnRNP proteins.

To gain insight into the significance of the RNA-binding spec-

ificity of hnRNP C1, we compared the sequence of the hnRNP C1 high affinity binding site with the sequences of known pre-mRNA processing signals (e.g. splice sites, polyadenylation signals, etc.). The vertebrate 3'-splice site consensus sequence (Mount, 1982; Ohshima and Gotoh, 1987) contains a loosely defined region, the polypyrimidine tract, that can potentially contain an hnRNP C1 high-affinity binding site. In fact, hnRNP C1/C2 has been shown to preferentially bind *in vitro* pre-mRNA fragments containing this region (Swanson and Dreyfuss, 1988b) and a monoclonal antibody to hnRNP C1/C2 inhibits *in vitro* pre-mRNA splicing (Choi *et al.*, 1986). Considering the affinity of hnRNP C1 for high-affinity binding sites and its estimated nuclear concentration (approximately 10  $\mu$ M (Kiledjian *et al.*, 1994)), it is likely that hnRNP C1 competes with other polypyrimidine tract binding proteins, such as U2 auxiliary factor (U2AF) (Zamore *et al.*, 1992) and hnRNP I/polypyrimidine tract binding protein (Patton *et al.*, 1991; Ghetti *et al.*, 1992; Bennet *et al.*, 1993), for binding to this site. The affinity of U2AF for different 3'-splice site regions ranges from  $10^{-8}$  to  $10^{-6}$  M (Zamore *et al.*, 1992). A potentially significant property of hnRNP C1 and its RBD derivative (as well as many other hnRNP proteins) is that they can promote the annealing of complementary RNA strands *in vitro* (Portman and Dreyfuss, 1994). This activity may reflect the ability of hnRNP C1 to influence interactions of other *trans*-acting factors with the RNA (Portman and Dreyfuss (1994) and references within) and could explain, at least in part, how multiple pre-mRNA binding factors with overlapping specificities can find high-affinity binding sites even when already bound to another protein. These observations support the suggestion that hnRNP C1 participates in pre-mRNA splicing.

**Protein Determinants of hnRNP C1 RNA-binding Specificity**—Amino acid sequence alignments of many RNP motif proteins suggested that the RBD of hnRNP C1/C2 extends from Lys<sup>8</sup> to Glu<sup>87</sup> (Dreyfuss *et al.*, 1988; Bandziulis *et al.*, 1989). To test this prediction, and to empirically delineate the RNA-binding domain of hnRNP C1, we tested the ability of *in vitro* produced full-length protein and carboxyl- and amino-terminal deletion mutants of hnRNP C1 to bind poly(rU) (Swanson and Dreyfuss, 1988a). This analysis (Fig. 2, A and B) revealed that carboxyl-terminal deletions to amino acid Lys<sub>84</sub> (K94) (Fig. 2,  $\Delta$ 94) do not significantly affect poly(rU) binding activity but removal of 9 more carboxyl-terminal amino acids abolishes binding (Fig. 2,  $\Delta$ 85). Deletion of the first 5 amino-terminal amino acids also abolishes binding (Fig. 2,  $\Delta$ 6/94). As expected, the 203 carboxyl-terminal amino acids of hnRNP C1 do not exhibit RNA binding activity (Fig. 2, HA- $\Delta$ 87/290). Thus, the amino-terminal 94 amino acids are necessary and sufficient for the poly(rU) binding activity of hnRNP C1 (see also Görlach *et al.* (1992)).

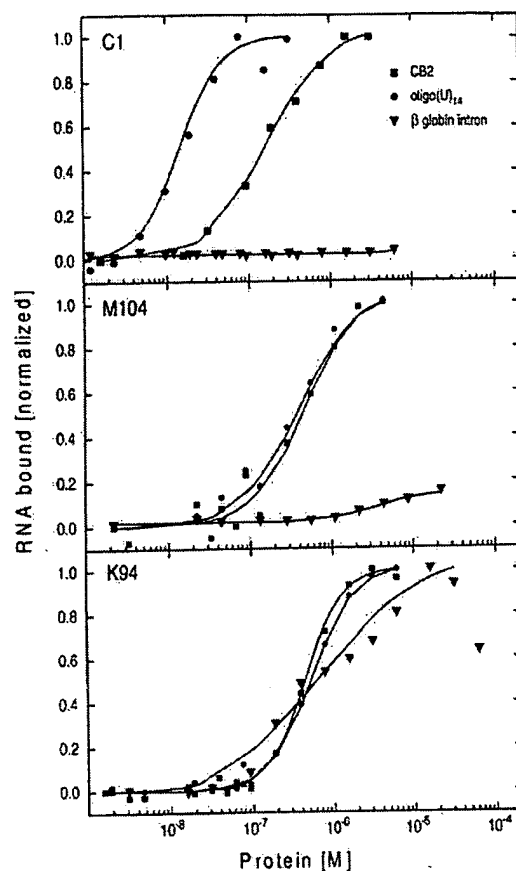
In order to more thoroughly characterize the RNA-binding properties of hnRNP C1 and its isolated RBD we performed gel mobility shift and filter-binding assays with a variety of RNA oligonucleotides containing high-affinity binding sites or other unrelated RNA sequences. For these experiments, hnRNP C1, K94, and a slightly larger fragment of hnRNP C1 spanning amino acids 1–104 (M104; see Fig. 2B,  $\Delta$ 104, and Fig. 4), were overexpressed in *E. coli*, purified, and used for the binding assays. Initial screening of the RNA binding activity of K94 and M104 toward a subset of *in vitro* selected oligonucleotides by gel-shift assays indicated that reducing the length of the RNA from 20 nucleotides to 14 or altering the sequences around the r(U)<sub>6</sub> minimal binding site did not significantly affect binding (data not shown). The affinities of all three polypeptides for CB2 (a 14-mer containing a central stretch of 5 U-residues) are similar (Fig. 3 and Table I). Full-length



**FIG. 2. Delineation of the minimal poly(rU)-binding domain of hnRNP C1.** *In vitro* transcribed and translated full-length hnRNP C1 and deletion mutants thereof were bound to poly(rU) linked to Sepharose beads. *In vitro* translated input material (lanes Total) along with protein bound to the poly(rU)-Sepharose was separated on a 12.5% NaDodecylSO<sub>4</sub>-polyacrylamide gel (lanes poly(rU)) and detected by fluorography. The lower half of the figure depicts the various deletion mutants of hnRNP C1 and their binding activity (+) or the lack thereof (-) toward poly(rU). The numbers indicate the position of the last or the first and the last, respectively, amino acid included in the construct with respect to the amino acid sequence of the full-length hnRNP C1 protein. The boxes indicate the canonical RBD of hnRNP C1 (Bandziulis *et al.*, 1989) with its RNP 1 and RNP 2 consensus sequences (stippled boxes). □, designates the hemagglutinin-tag sequence fused onto the Δ87/290 mutant to achieve sufficient translatability.

hnRNP C1 has only a 3-fold lower apparent  $K_d$  (170 nM) for CB2 than its RBD derivatives M104 and K94 (470 and 450 nM, respectively). However, there is a 30-fold difference in affinities between hnRNP C1 and its RBD constructs for the r(U)<sub>14</sub> oligoribonucleotide (14 nM (C1) versus 400 nM (M104) and 530 nM (K94)). From these experiments it appears that the binding activity of hnRNP C1 to an RNA with a single high-affinity binding site is faithfully reflected by its RBD derivatives (M104 or K94) and the observation that the RBD derivatives do not discriminate between r(U)<sub>5</sub> (i.e. CB2) and r(U)<sub>14</sub> suggests a cooperative involvement of COOH-terminal regions of hnRNP C1 in binding, even though this region does not exhibit RNA binding activity (Fig. 2). Finally, the affinity of K94 for oligo(rU)<sub>5</sub> in this assay ( $K_d$  = 5 μM; data not shown) is very similar to the  $K_d$  derived from our earlier NMR analysis of a K94-r(U)<sub>5</sub> complex (Görlach *et al.*, 1992).

A very interesting and unexpected result was obtained by comparing the binding activity of hnRNP C1 and its RBD derivatives to RNAs that do not contain high-affinity binding sites. Apparent dissociation constants were determined for the



**FIG. 3. Binding of RNA oligonucleotide substrates by hnRNP C1 and its RNA-binding domain.** Filter-binding assays were carried out and evaluated as described under "Materials and Methods" with (C1) hnRNP C1 and its RBD (M104) encompassing amino acid 1–104 or (K94) 1–94, respectively. RNA oligonucleotides were: ■, CB2 (AGUAUUUUUGUGGA); ●, r(U)<sub>14</sub>; ▼, 20-mer containing β-globin intron sequence (GAUCACUUGUGUCAACAG). Filter-bound RNA, normalized to saturation for each oligonucleotide is plotted as function of the respective protein concentration, corrected for the fraction of active protein. Curves were fitted to average data points of at least three independent binding experiments each.

β-globin intron-derived oligoribonucleotide and for oligo(rC)<sub>14</sub> (Fig. 3). Under the conditions of this assay, hnRNP C1 and M104 do not bind these RNAs even when present at micromolar concentrations, implying that the affinity of C1 and M104 for these ligands lies well above a  $K_d$  of 10 μM. In contrast, K94 bound both RNAs with an affinity approximately equal to its affinity for CB2 or r(U)<sub>14</sub> (Fig. 3; Table I). Apparently, K94 has lost its ability to discriminate between different RNA sequences and amino acids Ala<sup>95</sup> to Met<sup>104</sup> (Fig. 4) are mainly responsible for this activity. The lack of binding by hnRNP C1 and M104 to the β-globin intron-derived and (rC)<sub>14</sub> oligoribonucleotides is somewhat surprising because recent studies (Conway *et al.*, 1988; Huang *et al.*, 1994) found that hnRNP C1-C2 tetramers bind a wide variety of RNA sequences. Direct comparison of these experiments, however, is complicated by the different conditions used in each study (tetrameric hnRNP C1-C2 complexes with RNAs up to several thousand nucleotides in length (Conway *et al.*, 1988; Huang *et al.*, 1994) versus protein-excess binding experiments with short oligoribonucleotides (this study)). An interesting possibility is that the RNA binding activity of hnRNP C1 is affected by hnRNP C2.

The experiments described here, together with findings described for U1A and the U2B<sup>+</sup> (Scherly *et al.*, 1990a, 1990b; Jessen *et al.*, 1991), lead us to propose that the highly con-

MASNTNKT D PRSMNSRVFI GNLNTLVKK SDVEAIFSKY GKIVGCSVHK GFAFVQYVNE	60
RNRAAVAGE DGRMIAGQVL DINLAAEPKV NRKGAGVKRS AAEMVGSVTE HPSPSPPLISS	120
SFDLDYDFQR DYYDRMYSYP ARVPPPPPIA RAVVPSKRQR VSGNTSRRGK SGFNSKSGQR	180
GSSKSGKLKG DDLQAIKKEL TQIKQKVDSL LENLEKMEKE QSKQAVEMKN DKSEEEQSSS	240
SVKKDETNVK MESEGGADDS AEEGDLDDDD DNEDRGDDQL ELIKDDEKEA EEEGEDDRDSA	300
NGG	

FIG. 4. Amino acid sequence of hnRNP C1/C2. The amino acid sequence of the hnRNP C proteins C1 and C2 is shown. The minimal 94-amino acid RNA-binding domain (K94; Wittekind *et al.* (1992) and Görlich *et al.* (1992)) is boxed. The sequence specificity element (amino acids 95–104) of the RBD M104 is marked with the stippled box. The RNP 1 (underlined) and RNP 2 (hatched underlined) consensus sequences are shown. The 13 amino acid insert in hnRNP C2 (Burd *et al.*, 1989) is underlined with the stippled line.

served  $\beta\alpha\beta\beta\alpha$  core structure of the hnRNP C1 RBD constitutes a relatively nonsequence specific RNA-binding domain. In hnRNP C1, the major determinant of sequence-specific RNA binding resides immediately COOH-terminal to the minimal RBD. This idea is consistent with the results of experiments with other RNP motif proteins which also found that amino acids contiguous with, but outside of, RNP motif RBDs are important for RNA binding activity (Query *et al.*, 1989; Burd *et al.*, 1991; Scherly *et al.*, 1991; Cáceres and Krainer, 1993; Zuo and Manley, 1993). Furthermore, they support the idea that in the most variable regions of an RBD, *i.e.* the loops (Nagai *et al.*, 1990; Scherly *et al.*, 1990a, 1990b; Jessen *et al.*, 1991; Kenan *et al.* 1991) and contiguous NH<sub>2</sub> and COOH termini, provide much of the specificity of RNP motif RBDs. In this light, it is interesting to speculate that the 13-amino acid insert that distinguishes hnRNP C2 from hnRNP C1 (Fig. 4) and begins after Gly<sup>106</sup>, could directly influence RNA binding. In addition, hnRNP A2 and B1 differ only by a small peptide insert immediately adjacent to their first RNP motif RBD (Burd *et al.*, 1989). The model proposed here provides a rationale for how different members of this highly related class of proteins can specifically bind such diverse RNA ligands, despite the fact that they contain such highly conserved platform elements that mediate a large proportion of the RNA-protein interactions.

**Acknowledgments**—We thank Drs. F. Cobianchi and S. Riva for help with the construction of the overexpression plasmid for hnRNP C1, V. Apkarian for support with the oligonucleotide synthesis, and members of the laboratory for critically reading the manuscript.

#### REFERENCES

- Bandziulis, R. J., Swanson, M. S., and Dreyfuss, G. (1989) *Genes & Dev.* 3, 431–437.
- Bennet, M., Piñol-Roma, S., Staknis, D., Dreyfuss, G., and Reed R. (1993). *Mol. Cell. Biol.* 12, 3165–3175.
- Burd, C. G., and Dreyfuss, G. (1994) *EMBO J.* 13, 1197–1204.
- Burd, C. G., Swanson, M. S., Görlich, M., and Dreyfuss, G. (1989) *Proc. Natl. Acad. Sci. U. S. A.* 86, 9788–9792.
- Burd, C. G., Matunis, E. L., and Dreyfuss, G. (1991) *Mol. Cell. Biol.* 7, 3419–3424.
- Cáceres, J. F., and Krainer, A. R. (1993) *EMBO J.* 12, 4715–4726.
- Choi, Y. D., and Dreyfuss, G. (1984) *J. Cell Biol.* 99, 1997–2004.
- Choi, Y. D., Grabowski, P. J., Sharp, P. A., and Dreyfuss, G. (1986) *Science* 231, 1534–1539.
- Cobianchi, F., Karpel, R. L., Williams, K. R., Notario, V., and Wilson, S. H. (1988) *J. Biol. Chem.* 263, 1063–1071.
- Conway, G., Wooley, J., Bibring, T., and LeStourgeon, W. M. (1988) *Mol. Cell. Biol.* 8, 2884–2895.
- Crowl, R., Seamans, C., Lomedico, P., and Andrew, S. (1985) *Gene (Amst.)* 38, 31–38.
- Dreyfuss, G. (1986) *Annu. Rev. Cell Biol.* 2, 459–498.
- Dreyfuss, G., Choi, Y. D., and Adam, S. A. (1984) *Mol. Cell. Biol.* 4, 1104–1114.
- Dreyfuss, G., Swanson, M. S., and Piñol-Roma, S. (1988) *Trends Biochem. Sci.* 13, 86–91.
- Dreyfuss, G., Matunis, M. J., Piñol-Roma, S., and Burd, C. G. (1993) *Annu. Rev. Biochem.* 62, 289–321.
- Ghetti, A., Piñol-Roma, S., Michael, W. M., Morandi, C., and Dreyfuss, G. (1992) *Nucleic Acids Res.* 20, 3671–3678.
- Görlich, M., Burd, C. G., Portman, D. S., and Dreyfuss, G. (1993) *Mol. Biol. Reports* 18, 73–78.
- Görlich, M., Wittekind, M., Beckman, R. A., Mueller, L., and Dreyfuss, G. (1992) *EMBO J.* 11, 3289–3295.
- Hoffman, D. W., Query, C. C., Golden, B. L., White, S. W., and Keene, J. D. (1991) *Proc. Natl. Acad. Sci. U. S. A.* 88, 2495–2499.
- Huang, M., Rech, J. E., Northington, S. J., Flicker, P. F., Mayeda, A., Krainer, A. R., and LeStourgeon, W. M. (1994) *Mol. Cell. Biol.* 14, 518–533.
- Jessen, T.-H., Oubridge, C., Teo, C. H., Pritchard, C., and Nagai, K. (1991) *EMBO J.* 10, 3447–3456.
- Kelly, R. C., Jensen, D. E., and von Hippel, P. H. (1976) *J. Biol. Chem.* 251, 7240–7250.
- Kenan, D. J., Query, C. C., and Keene J. D. (1991) *Trends Biochem. Sci.* 16, 214–220.
- Kiledjian M., Burd, C. G., Görlich, M., Portman, D. S., and Dreyfuss, G. (1994) in *Structure and Function of hnRNP Proteins in RNA-protein Interactions-Frontiers in Molecular Biology* (Nagai, K., and Mattaj, I., eds) Oxford University Press, Oxford, UK, in press.
- Lutz-Freyermuth, C., Query, C. C., and Keene, J. D. (1990) *Proc. Natl. Acad. Sci. U. S. A.* 87, 6393–6397.
- Mattaj, I. W. (1989) *Cell.* 57, 1–3.
- Matunis, M. J., Matunis, E. L., and Dreyfuss G. (1992) *J. Cell Biol.* 116, 245–255.
- Matunis, E. L., Matunis, M. J., and Dreyfuss G. (1993) *J. Cell Biol.* 121, 219–228.
- Merrill, B. M., Stone, K. L., Cobianchi, F., Wilson, S. H., and Williams, K. R. (1988) *J. Biol. Chem.* 263, 3307–3313.
- Mount, S. M. (1982) *Nucleic Acids Res.* 10, 459–472.
- Nagai, K., Oubridge, C., Jessen, T. H., Li, J., and Evans, P. R. (1990) *Nature* 346, 515–520.
- Nietfeld, W., Mentzel, H., and Pieler, T. (1990) *EMBO J.* 9, 3699–3705.
- Ohshima, Y., and Gotoh, Y. (1987) *J. Mol. Biol.* 195, 247–259.
- Patton, J. G., Meyer, S. A., Tempst, P., and Nadal-Ginard, B. (1991) *Genes & Dev.* 5, 1237–1251.
- Piñol-Roma, S., Choi, Y. D., Matunis, M. J., and Dreyfuss, G. (1988) *Genes & Dev.* 2, 215–227.
- Portman, D., and Dreyfuss, G. (1994) *EMBO J.* 13, 213–221.
- Query, C. C., Bentley, R. C., and Keene, J. D. (1989) *Cell* 57, 89–101.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Scherly, D., Boelens, W., van Venrooij, W. J., Dathan, N. A., Hamm, J., and Mattaj, I. W. (1989) *EMBO J.* 8, 4163–4170.
- Scherly, D., Boelens, W., Dathan, N. A., van Venrooij, W. J., and Mattaj, I. W. (1990a) *Nature* 345, 502–506.
- Scherly, D., Dathan, N. A., Boelens, W., van Venrooij, W. J., and Mattaj, I. W. (1990b) *EMBO J.* 9, 3675–3681.
- Scherly, D., Kambach, C., Boelens, W., van Venrooij, W. J., and Mattaj, I. W. (1991) *J. Mol. Biol.* 219, 577–584.
- Schwemmler, M., Görlich, M., Bader, M., Sarre, Th. F., and Hilse, K. (1989) *FEBS Lett.* 251, 117–120.
- Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Methods Enzymol.* 185, 60–89.
- Surovy, C. S., van Santen, V. L., Scheib-Wixted, S. M., and Spritz, R. A. (1989) *Mol. Cell. Biol.* 9, 4179–4186.
- Swanson, M. S., and Dreyfuss, G. (1988a) *Mol. Cell. Biol.* 8, 2237–2241.
- Swanson, M. S., and Dreyfuss, G. (1988b) *EMBO J.* 7, 3519–3529.
- Swanson, M. S., Nakagawa, T. Y., LeVan, K., and Dreyfuss, G. (1987) *Mol. Cell. Biol.* 7, 1731–1739.
- Tuerk, C., and Gold, L. (1990) *Science* 249, 505–510.
- Wilusz, J., and Shenk, T. (1990) *Mol. Cell. Biol.* 10, 6397–6407.
- Wilusz, J., Feig, D. L., and Shenk, T. (1988) *Mol. Cell. Biol.* 8, 4477–4483.
- Witherell, G. W., and Uhlenbeck, O. C. (1989) *Biochemistry* 28, 71–76.
- Wittekind, M., Görlich, M., Friedrichs, M., Dreyfuss, G., and Mueller, L. (1992) *Biochemistry* 31, 6254–6265.
- Zamore, P. D., Patton, J. G., and Green, M. R. (1992) *Nature* 355, 609–614.
- Zuo, P., and Manley, J. L. (1993) *EMBO J.* 12, 4727–4737.

**Experimental Cell Research**

Volume 211, Issue 2 , April 1994, Pages 400-407

**The mRNA Poly(A)-Binding Protein: Localization, Abundance, and RNA-Binding Specificity****Matthias Görlach, Christopher G. Burd and Gideon Dreyfuss**

Howard Hughes Medical Institute and Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-6148

Available online 29 April 2002.

**Abstract**

The poly(A)-binding protein (PABP) binds to the messenger (mRNA) 3'-poly(A) tail found on most eukaryotic mRNAs and together with the poly(A) tail has been implicated in governing the stability and the translation of mRNA. In order to further understand the role of the PABP in these processes, we have undertaken a detailed analysis of the cellular localization, the abundance, and the RNA-binding properties of the human PABP (hPABP). We raised monoclonal antibodies against the 70-kDa hPABP and confocal immunofluorescence microscopy with these antibodies reveals that it is localized exclusively to the cytoplasm. The hPABP exhibits a very low turnover rate in these cells and quantitative immunoblotting experiments demonstrated that growing HeLa cells contain a surprisingly high number of approximately  $8 \times 10^6$  PABP molecules per cell, which corresponds to an intracellular concentration of about  $4 \mu M$ . In an *in vitro* selection/amplification assay from random sequence oligonucleotide pools the hPABP selects oligo(rA)-rich sequences and it binds oligo(rA)<sub>25</sub> with an apparent  $K_d$  of 7 nM. The hPABP binds to unrelated RNA sequences with an about 100-fold lower affinity ( $K_d \geq 0.5 \mu M$ ). The abundance of the hPABP indicates that there is an approximately three-fold excess of the protein over binding sites on cytoplasmic poly(A). This excess and the high concentration of the hPABP, which is three orders of magnitude above its  $K_d$  for oligo(rA)<sub>25</sub>, suggest that the hPABP may bind to additional, lower affinity binding sites *in vivo*.

# Wheat Germ Poly(A) Binding Protein Enhances the Binding Affinity of Eukaryotic Initiation Factor 4F and (iso)4F for Cap Analogues<sup>†</sup>

Chin-Chuan Wei, M. Luisa Balasta, Jianhua Ren, and Dixie J. Goss\*

Department of Chemistry of Hunter College of the City University of New York, 695 Park Avenue, New York, New York 10021

Received October 3, 1997; Revised Manuscript Received December 3, 1997

**ABSTRACT:** Most eukaryotic mRNAs contain a 5' cap (m<sup>7</sup>GpppX) and a 3' poly(A) tail to increase synergistically the translational efficiency. Recently, the poly(A) binding protein (PABP) and cap-binding protein, eIF-4F, were found to interact [Le et al. (1997) *J. Biol. Chem.* 272, 16247–16255; Tarun and Sachs (1996) *EMBO J.* 15, 7168–7177]. These data suggest that PABP may exert its effect on translational efficiency either by increasing the formation of initiation factor–mRNA complex or by enhancing ribosome recycling. To investigate the functional consequences of these interactions, the fluorescent cap analogue, ant-m<sup>7</sup>GTP, which is an environmentally sensitive fluorescent probe [Ren and Goss (1996) *Nucleic Acids Res.* 24, 3629–3634] was used to investigate the cap-binding affinity. Our data show that the binding of eIF-(iso)4F or eIF-4F to cap analogue enhanced their binding affinity toward PABP approximately 40-fold. Similarly, the eIF-4F/PABP or eIF-(iso)4F/PABP complexes show a 40-fold enhancement of cap analogue binding as compared to eIF-4F or eIF-(iso)4F alone. At least part of the enhancement of the translational initiation by PABP can be accounted for by direct changes in cap-binding affinity. The interactions of these components also suggest a mechanism whereby the poly(A) tail is brought into close proximity with m<sup>7</sup>G cap. This effect was examined by fluorescence energy transfer, and it was determined that the PABP/eIF-4F complex could bind both poly(A) and 5' cap simultaneously.

In most eukaryotes, mRNA is required to have both a 5' cap (m<sup>7</sup>GpppX)<sup>1</sup> and a poly(A) tail for efficient translation and message stability (1–6). These two elements act synergistically to increase translational efficiency, and recent evidence suggests that they communicate during translation (7–9). The cap serves as the binding site for initiation factors eIF-4F and eIF-(iso)4F, an isozyme form of eIF-4F present in higher plants. The small subunit of eIF-4F (eIF-4E) recognizes the cap structure. eIF-4F interacts with the poly(A) binding protein (PABP) through the 4G subunit, the larger subunit of eIF-4F or eIF-(iso)4F. The cap-associated proteins have a very high binding affinity (nM) for PABP in the absence of poly(A) in the wheat germ system (9) but require poly(A) in yeast (8). The synergistic effects on translational efficiency may result directly from increased cap affinity of the complex, from efficient recycling of ribosomes, or from a combination of both of these mechanisms. Earlier work (9) has shown that PABP interaction with the cap-binding proteins enhanced the affinity for poly(A). This leaves unanswered the question of how interactions at the 3' end of the mRNA increase translational efficiency. Does PABP also increase the affinity of initiation factors for the cap at the 5' terminus of mRNA, and if so, is the protein complex of PABP and eIFs capable of binding both the cap and poly(A) simultaneously? Simultaneous binding

of both ends of the mRNA is necessary if looping of the mRNA occurs and plays a role in enhancement of translation.

To understand the mechanism of translational enhancement, we have investigated the binary and ternary interactions among the various proteins, m<sup>7</sup>G cap analogue, and poly(A) using fluorescence spectroscopy. The data analysis quantitates the interactions of cap-associated proteins and PABP. These quantitative results have allowed us to determine the binding affinity of the cap–protein complex for PABP and the effects of PABP on the cap-binding affinity of eIF-4F and eIF-(iso)4F. PABP increased the cap-binding affinity of both proteins by approximately 40-fold. Furthermore, fluorescence energy transfer experiments demonstrate that the protein complex is capable of binding both cap and poly(A) simultaneously. These results suggest the possibility of RNA looping as a means of translational enhancement and that the synergistic effect of PABP on protein synthesis can be accounted for at least partially by a direct effect on cap affinity of the initiation factors.

## EXPERIMENTAL PROCEDURES

**Materials.** m<sup>7</sup>GTP was purchased from Sigma (St. Louis, MO). Nuclease P<sub>1</sub> and poly(A) were purchased from Pharmacia Biotech (Uppsala, Sweden). The synthesis of ant-m<sup>7</sup>GTP was carried out as described previously (10). The preparation of PABP followed the procedure of Le et al. (9) and Yang and Hunt (11). The eIF-4F and eIF-(iso)4F were prepared as described elsewhere (12, 13). The concentration of ant-m<sup>7</sup>GTP was determined spectrophotometrically using an absorption coefficient of  $\epsilon_{332} = 4600 \text{ M}^{-1} \text{ cm}^{-1}$ . Poly(A)<sub>30</sub> was prepared by heating 10 mg of poly(A) in the presence of 21 units of Nuclease P<sub>1</sub> at 37 °C, followed by phenol extraction, dialysis, and electrophoresis in 15%

<sup>†</sup> This work was supported by grants from the National Science Foundation (GER-9023681 and MCB-9722907 to D.J.G.).

\* To whom correspondence should be addressed. Tel: (212) 772-5383, fax: (212) 772-5332.

<sup>1</sup> Abbreviations: eIF, eukaryotic initiation factor; m<sup>7</sup>G, 7-methylguanosine; ant-m<sup>7</sup>GTP or cap\*, anthraniloyl 7-methylguanosine triphosphate; Tris, tris(hydroxymethyl)aminomethane; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; BSA, bovine serum albumin; kDa, kilodalton(s).

polyacrylamide gel containing 6.0 M urea and a running buffer containing 0.09 M Tris-borate and 0.002 M EDTA. The gel was examined under UV light. The desired length of poly(A) was compared with oligonucleotide markers. The poly(A)<sub>30</sub> was cut from the gel, and this gel slice was crushed and eluted with buffer containing 10 mM Tris, pH 8.0, and 0.1 mM EDTA. The 3'-labeling of poly(A)<sub>30</sub> followed the procedure of Odom et al. (14) with some modifications. All reactions were carried out in the dark. Oligonucleotide was oxidized at the 3' end by incubation with 0.09 M sodium periodate in 0.1 M sodium acetate, pH = 6.0, for 1 h at 37 °C. The reaction was terminated by the addition of KCl. The KIO<sub>4</sub> precipitate was removed by centrifugation at 5000g at 4 °C for 10 min. The supernatant was dialyzed against 0.1 M sodium acetate, pH 6.0, at 4 °C. The sample was incubated with 1 mM fluorescein thiosemicarbazide at 37 °C for 2 h. At the end of the incubation, the sample was purified by Sephadex G-25 column chromatography. All samples were dialyzed against buffer A (25 mM Tris, 100 mM KOAc, 1 mM CaCl<sub>2</sub>, 1 mM DDT, 1 mM EDTA, 1 mM MgCl<sub>2</sub>, pH = 7.6) and passed through a 0.22-μm filter (Millipore, MA) before the spectroscopy measurements were performed. Protein concentrations were estimated by the method of Bradford (15) using a Bio-Rad protein assay reagent (Bio-Rad Laboratories, CA).

**Spectroscopy Measurements.** Absorbance measurements were obtained using a Cary-3 double-beam UV/VIS spectrophotometer. Fluorescence spectra were recorded on a Spex Fluorolog τ<sub>2</sub> spectrofluorometer equipped with excitation and emission polarizers. All measurements were performed at 20 °C. For all equilibrium measurements, at least three titrations were performed, and for fluorescence energy transfer, three reproducible results were obtained. The variation in values for equilibrium constants from different titrations were within 6%.

For measurements of protein-protein interactions, complex formation was studied by monitoring changes in intrinsic protein fluorescence. To maximize the signal, the excitation wavelength was chosen to be 265 nm and emission signal was detected with a cuton filter (50% transmission at 305 nm). Since no emission monochromator was used, the fluorescence intensity change due to fluorescence polarization did not need to be considered. The signal was corrected for light scattering by measuring the buffer under the same conditions. The excitation band-pass was chosen to eliminate photobleaching, and the linearity of fluorescence intensity with protein concentration was determined. For each data point, three samples were prepared. The fluorescence intensity of a solution containing 100 nM PABP was measured. A second sample with a specific amount of cap-binding protein was also measured, and the corrected intensities of the two samples were summed together (*F<sub>s</sub>*). A third sample containing the same amount of PABP and cap-binding protein mixed together was incubated at 20 °C for 20 min, and the corrected fluorescence intensity for this complex was obtained (*F<sub>c</sub>*). The difference in fluorescence intensity related to the complex was defined as Δ*F* = *F<sub>c</sub>* - *F<sub>s</sub>*. The details of the fitting are described elsewhere (16, 17).

For protein-nucleic acid interactions, the excitation wavelength for ant-m<sup>7</sup>GTP was 332 nm and emission was monitored at 420 nm. The excitation slits were chosen to

avoid photobleaching, and the absorbance of the sample at the excitation wavelength was less than 0.02 to minimize the inner-filter effect. Emission spectra were corrected for wavelength-dependent lamp intensity and monochromator sensitivities. Steady-state fluorescence anisotropy was measured using an L-format detection configuration. The excitation band-pass was 8.5 nm, and the emission band-pass was 17.0 nm. All samples were incubated at least 15 min at 20 °C before data were collected.

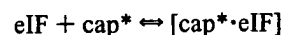
Phase-demodulation lifetime measurements were carried out using a 0.1% (w/w) glycogen suspension in water with a reference lifetime of 0.0 ns. The emission polarizer was placed at a magic angle (54°) to correct for the different sensitivities of the polarized light on the photomultiplier tube and the monochromator.

To determine if energy transfer occurred between the cap-binding analogue and poly(A) when bound to the protein complex, 150 nM ant-m<sup>7</sup>GTP, 50 nM eIF-4F, 50 nM PABP, and 50 nM 3'-fluorescein-poly(A)<sub>30</sub> were mixed together for complex formation with the absorbance at the excitation wavelength of less than 0.02. The fluorescence intensity (*F<sub>D,A</sub>*) from 380 to 600 nm was recorded. Unlabeled m<sup>7</sup>GTP and poly(A)<sub>30</sub> were used as controls since they have almost identical binding affinity as compared with ant-m<sup>7</sup>GTP and 3'-fluorescein-poly(A)<sub>30</sub>, respectively. The fluorescence signal from donor (*F<sub>D</sub>*) in the absence of acceptor was determined using 150 nM ant-m<sup>7</sup>GTP, 50 nM eIF-4F, 50 nM PABP, and 50 nM poly(A)<sub>30</sub>. The fluorescence signal from acceptor (*F<sub>A</sub>*) at this excitation wavelength was obtained using 150 nM m<sup>7</sup>GTP, 50 nM eIF-4F, 50 nM PABP, and 50 nM 3'-fluorescein-poly(A)<sub>30</sub>. Comparison of the spectra obtained from the mixture of both donor and acceptor with the spectra obtained from the sum of individual fluorescence intensities of controls allowed detection of fluorescence energy transfer. The spectral overlap integral, *J*, of the emission spectra of the donor and the absorption spectra of the acceptor was calculated according to

$$J = \frac{\int F_D(\lambda) \epsilon_A(\lambda) \lambda^2 d\lambda}{\int F_D(\lambda) \lambda^{-2} d\lambda}$$

where *F<sub>D</sub>*(λ) is the fluorescence spectrum of the donor and ε<sub>A</sub>(λ) is the molar absorptivity of the acceptor on a wavelength scale λ. The Förster distance, *R<sub>0</sub>*, between two dyes was not calculated because κ<sup>2</sup>, a function of the relative orientation of the dyes, could not be measured directly and the orientation was unlikely to be random.

**Data Analysis.** The equilibrium constant (*K<sub>a</sub>*) for the binding of ant-m<sup>7</sup>GTP with eIFs



is defined by

$$K_a = \frac{[\text{cap}^* \cdot \text{eIF}]}{[\text{cap}^*][\text{eIF}]} \quad (1)$$

where [eIF], [cap\*], and [cap\*·eIF] are the equilibrium concentrations of the unbound protein, ant-m<sup>7</sup>GTP, and ant-m<sup>7</sup>GTP/protein complex, respectively.

The fluorescence anisotropy, *r*, is defined as

$$r = \frac{I_{VV} - GI_{VH}}{I_{VV} + 2GI_{VH}} \quad (2)$$

where  $G$  is a factor that accounts for the polarization bias of the detection system:

$$G = \frac{I_{HV}}{I_{HH}}$$

where  $I_{HV}$  and  $I_{HH}$  are fluorescence intensity measured with horizontal excitation polarization and with the emission polarizer aligned either vertically or horizontally, respectively.

In a mixture of eIF and ant-m<sup>7</sup>GTP, the average anisotropy is related to the fraction of the total fluorophore that is bound ( $f_b$ ) by

$$f_b = \frac{r - r_f}{(r_b - r)R + (r - r_f)} \quad (3)$$

where  $R = Q_b/Q_f$  is the ratio of quantum yield of the bound and the free ant-m<sup>7</sup>GTP.  $r$ ,  $r_f$ , and  $r_b$  are the anisotropy values of the mixture, the free ant-m<sup>7</sup>GTP, and the totally bound ant-m<sup>7</sup>GTP. It can be shown that the measured anisotropy is related to the total concentration of eIF by

$$r = \frac{ar_bR - ar_f + r_f}{1 + a(R - 1)} \quad (4)$$

where

$$a = \frac{(K[\text{cap}^*]_T + K[\text{eIF}]_T + 1) - \sqrt{(K[\text{cap}^*]_T + K[\text{eIF}]_T + 1)^2 - 4K^2[\text{cap}^*]_T[\text{eIF}]_T}}{2K[\text{cap}^*]_T}$$

$[\text{cap}^*]_T$ ,  $[\text{eIF}]_T$ , and  $K$  were the total concentration of ant-m<sup>7</sup>GTP, protein, and the equilibrium association constant, respectively.

Multifrequency phase and modulation data were analyzed by fitting to a sum of exponentials using the Globals Unlimited program (Urbana, IL).

The measurement of relative quantum yield,  $Q$ , of fluorophore was calculated by comparison to a standard whose quantum yield is known (18):

$$Q_x = \frac{I_x Q_s A_s}{I_s A_x}$$

where  $x$  represents the unknown species,  $s$  represents the standard,  $I$  is the integrated amount of fluorescence, and  $A$  is the absorbance at the excitation wavelength. Disodium fluorescein in 0.01 N NaOH,  $Q = 0.92$  (19), was used as a standard.

## RESULTS

Recently, the binding affinities of cap-associated proteins, eIF-4F, eIF-(iso)4F, and eIF-4B, for PABP were determined (9). The data showed the dissociation constants ( $K_d$ ) were less than 40 nM for eIF-4F and eIF-(iso)4F. For eIF-4B, the  $K_d$  was approximately 15 nM. To determine the binding affinities quantitatively, more dilute protein solutions were used. A cuton filter (50% transmission at 305 nm) was used

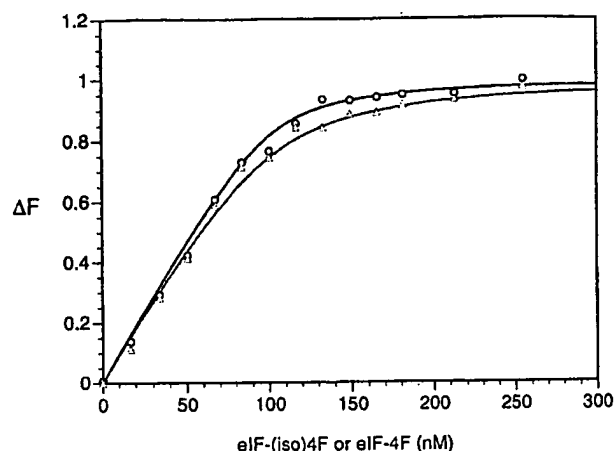


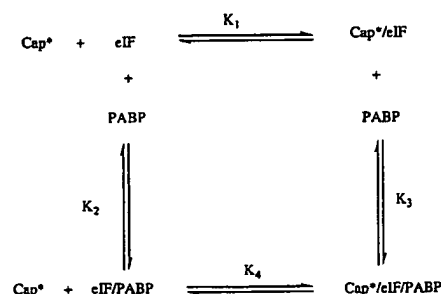
FIGURE 1: Solution of 100 nM PABP was titrated with purified eIF-(iso)4F (circles) and eIF-4F (triangles). The difference between the fluorescence intensities for the combined proteins and the sum of fluorescence intensities of the individual proteins was plotted as a function of the concentration of eIF-(iso)4F or eIF-4F.

Table 1

interactions	$K_1^a$ ( $\mu\text{M}$ )	$K_2$ (nM)	$K_3$ (nM)	$K_4$ ( $\mu\text{M}$ )
ant-m <sup>7</sup> GTP/eIF-(iso)4F/PABP	$8.93 \pm 1.04$	$4.3 \pm 1.9$	$0.099 \pm 0.045$	$0.21 \pm 0.09$
ant-m <sup>7</sup> GTP/eIF-4F/PABP	$4.69 \pm 0.33$	$9.1 \pm 4.1$	$0.23 \pm 0.05$	$0.12 \pm 0.06$

<sup>a</sup>  $K$  represents the dissociation constant. See Scheme 1 for detail of interactions.

Scheme 1



in place of an emission monochromator to monitor the fluorescence signal. The titration curve in Figure 1 shows the difference in fluorescence intensity between the protein-protein complex and the sum of individual fluorescence intensities. The binding curve extrapolates to 1:1 stoichiometry for the two proteins. As a further control, BSA was used to test for nonspecific binding. The fluorescence signal did not show significant change upon addition of BSA (data not shown). The  $K_d$  values obtained from data fitting of three curves for the two cap-binding proteins, eIF-4F and eIF-(iso)4F, interacting with PABP are shown in Table 1 ( $K_2$  in Table 1 and Scheme 1). The initiation factor, eIF-4A, which does not bind to the cap region showed no interaction with PABP as reported earlier (9).

The fluorescence intensity and maximum emission wavelength of the cap analogue, ant-m<sup>7</sup>GTP, increased when bound to eIF-4F or eIF-(iso)4F (10). The corrected fluorescence spectra for this free cap analogue showed an excitation maximum at 332 nm and emission maximum at 420 nm (10). The anisotropy ranged from 0.03 for unbound

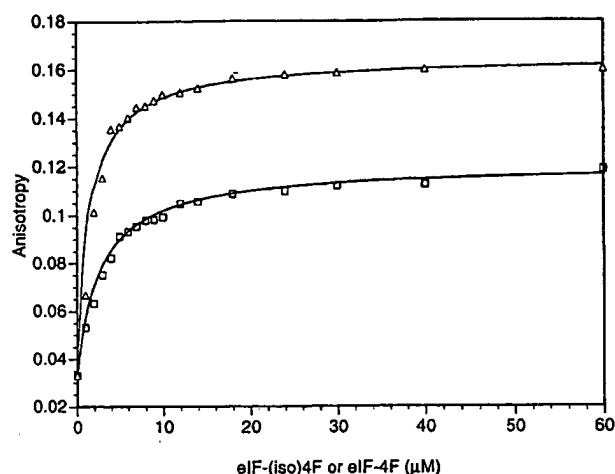


FIGURE 2: Titration of ant-m<sup>7</sup>GTP with eIF-4F (triangles) and eIF-(iso)4F (squares). Titrations were performed in buffer A. Concentration of ant-m<sup>7</sup>GTP was 0.3 μM. Complex formation was monitored by steady-state anisotropy. The solid line presents the best fit of the data to eq 4.

fluorophore to 0.12 for ant-m<sup>7</sup>GTP/eIF-(iso)4F (Figure 2). The ant-m<sup>7</sup>GTP/eIF-4F complex had an anisotropy value of 0.16. By fitting the data according to eq 4, the  $K_d$  values ( $K_1$  in Table 1 and Scheme 1) were calculated to be  $8.93 \pm 1.04$  μM for ant-m<sup>7</sup>GTP/eIF-(iso)4F and  $4.69 \pm 0.07$  μM for ant-m<sup>7</sup>GTP/eIF-4F (Figure 2). These data are similar to those obtained for m<sup>7</sup>G<sub>ppp</sub>G binding (20), which is consistent with the fact that most nucleotide-binding proteins are sensitive to structural variations in the purine ring, but modifications of the ribose moiety have little effect on binding affinity (21–23). The characteristic anisotropy of the ant-m<sup>7</sup>GTP/eIF-4F complex was higher than that of the ant-m<sup>7</sup>GTP/eIF-(iso)4F complex because the eIF-4F has the almost twice the molecular weight of eIF-(iso)4F (246 kDa as compared to 114 kDa) while the fluorescence lifetimes of the complexes were similar. The observed steady-state anisotropy depends on the correlation time ( $t_c$ ) of the fluorophore and its lifetime (24). Free ant-m<sup>7</sup>GTP has a fluorescence excited-state lifetime of 2.0 ns (Figure 3). Lifetime measurements on the ant-m<sup>7</sup>GTP/eIF-(iso)4F showed a longer lifetime of 6.8 ns attributable to the protein complex. The ant-m<sup>7</sup>GTP/eIF-4F complex has a shorter lifetime of 5.0 ns. The fluorescence enhancement and longer lifetime of the protein-bound form of ant-m<sup>7</sup>GTP may be due to the formation of hydrogen bond(s), which stabilize its charge-transfer excited state, or the cap-binding pocket may be a more hydrophobic region (25). The recent X-ray crystal structure of murine eIF-4E (26) indicates that 7-methyl G recognition is mediated by the guanosine base interaction between two conserved tryptophans plus formation of three hydrogen bonds and a van der Waals interaction between its N-7 methyl group and a third conserved tryptophan. The ribose of the cap is located in a hydrophobic region, which may allow the nonprotonated chromophore of the ant-m<sup>7</sup>GTP to enter the binding pocket. Stabilization of the polar excited state would result in a longer lifetime.

It has been previously reported (7) that eIF-(iso)4F, eIF-4F, and eIF-4B bound to the mRNA poly(A) tail. The size of the binding sites of eIF-4F and eIF-4B were estimated to be 25 bases in a buffer containing 150 mM KCl (27). From the packing density measurement, PABP was also found to

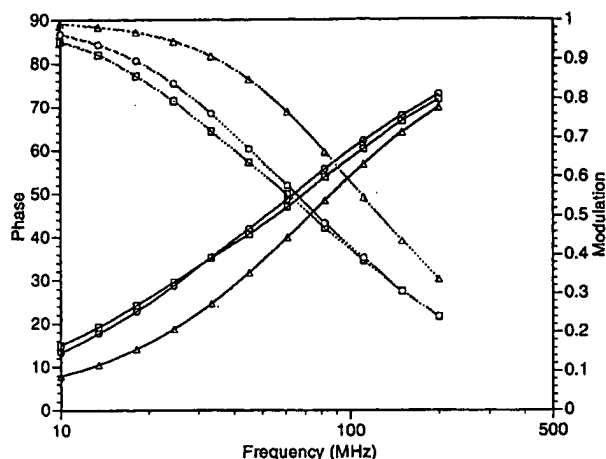


FIGURE 3: Phase and modulation values as a function of frequency for ant-m<sup>7</sup>GTP bound to eIF-(iso)4F, ant-m<sup>7</sup>GTP alone (triangles); The solid line was fitted as a one-exponential decay with a lifetime of 2.0 ns. ant-m<sup>7</sup>GTP and eIF-(iso)4F (squares): The solid line corresponds to a two-exponential component fit of the data with lifetimes  $\tau_1 = 1.98$  ns and  $\tau_2 = 6.93$  ns having a fractional amplitude  $f_1 = 0.513$  and  $f_2 = 0.487$ . ant-m<sup>7</sup>GTP and eIF-4F (circles): The solid line represents  $\tau_1 = 2.01$  ns and  $\tau_2 = 5.04$  ns having a fraction amplitude  $f_1 = 0.404$  and  $f_2 = 0.596$ .

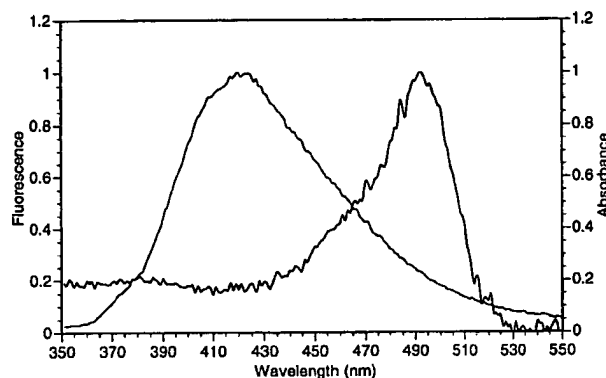


FIGURE 4: Spectral overlap of absorption spectrum (thick line) of 3'-fluorescein-poly(A)<sub>30</sub> and emission spectrum (thin line) of ant-m<sup>7</sup>GTP in the buffer A. For presentation, absorbance and fluorescence were normalized to have the same maximum values.

bind 25 A residues per molecule (28). To determine whether the cap and poly(A) tail were bound to separate sites on this protein complex, energy transfer experiments were performed. The fluorescence probe, fluorescein, was attached to the 3' end of poly(A)<sub>30</sub>. The spectral overlap between fluorescence emission of ant-m<sup>7</sup>GTP and absorbance of fluorescein labeled poly(A) (Figure 4) allowed fluorescence energy transfer to be monitored. Titration of 3'-fluorescein-poly(A)<sub>30</sub> with eIF-4F and PABP showed no significant fluorescence intensity changes, but an anisotropy change was observed indicating formation of a nucleic acid–protein complex. As a control, m<sup>7</sup>GTP and unlabeled poly(A)<sub>30</sub> were used. Figure 5 shows the spectrum obtained from the complex, ant-m<sup>7</sup>GTP/eIF-4F/PABP/3'-fluorescein-poly(A)<sub>30</sub> and the spectrum from the sum of the spectra for m<sup>7</sup>GTP/eIF-4F/PABP/3'-fluorescein-poly(A)<sub>30</sub> and ant-m<sup>7</sup>GTP/eIF-4F/PABP/poly(A)<sub>30</sub>. Essentially the same results were obtained with three different experiments using two different PABP preparations. The fluorescence quenching at 420 nm and enhancement at 510 nm demonstrated that eIF-4F/PABP

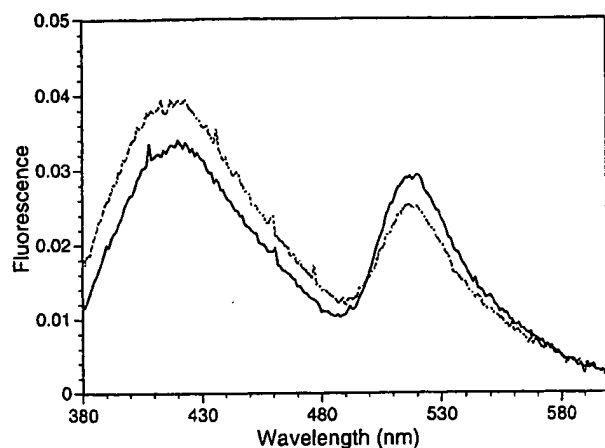


FIGURE 5: Fluorescence energy transfer. The solid line is the measured fluorescence intensity of ant-m<sup>7</sup>GTP/eIF-4F/PABP/3'-fluorescein-poly(A)<sub>30</sub>. The broken line is the sum of the fluorescence spectra for m<sup>7</sup>GTP/eIF-4F/PABP/3'-fluorescein-poly(A)<sub>30</sub> and ant-m<sup>7</sup>GTP/eIF-4F/PABP/poly(A)<sub>30</sub>. The increase in acceptor fluorescence at 510 nm and corresponding decrease in fluorescence at 420 nm indicates energy transfer between the cap and poly(A) tail. The concentration used for ant-m<sup>7</sup>GTP (m<sup>7</sup>GTP), eIF-4F, PABP, and 3'-fluorescein-poly(A)<sub>30</sub> (poly(A)<sub>30</sub>) were 150, 50, 50, and 50 nM, respectively.

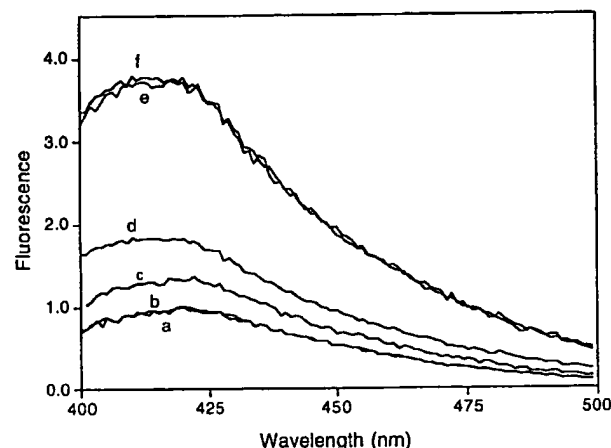


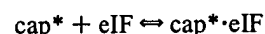
FIGURE 6: Corrected fluorescence emission spectra of 25 nM ant-m<sup>7</sup>GTP (pH 7.5) (a) ant-m<sup>7</sup>GTP alone, (b) +0.4 μM PABP, (c) +0.02 μM eIF-(iso)4F, (d) +0.02 μM eIF-(iso)4F and 0.01 μM PABP, (e) +1.5 μM eIF-(iso)4F, and (f) +1.5 μM eIF-(iso)4F and 25 μM PABP. All samples were excited at 332 nm.

bound both fluorescent cap analogue and fluorescent poly(A) simultaneously. At these concentrations, the two proteins will be present almost entirely as a complex. The cap analogue, however, will have some fluorophore unbound. This may partially explain the efficiency of the energy transfer.

For ternary ant-m<sup>7</sup>GTP/eIF-(iso)4F/PABP and ant-m<sup>7</sup>GTP/eIF-4F/PABP studies, steady-state fluorescence spectra were examined. Figure 6 shows the emission spectra of different ant-m<sup>7</sup>GTP complexes excited at 332 nm. The ant-m<sup>7</sup>GTP, when excited at 332 nm, shows a maximum fluorescence intensity at 420 nm. Upon addition of PABP, no fluorescence intensity change was observed, indicating that no interaction between this cap analogue and PABP occurred. This result was confirmed using anisotropy and lifetime measurements (data not shown), which also do not show any change upon addition of PABP. Fluorescence intensity was

enhanced and the emission wavelength maximum shifted when eIF-(iso)4F was added to the cap analogue. Under conditions where concentrations of eIF-(iso)4F and PABP are larger than the concentration of cap analogue, no significant fluorescence intensity difference between cap analog/eIF complex and cap analog/eIFs/PABP complex was observed. The lifetime analysis on this ternary system could be fitted as two-exponential decay with the same results as ant-m<sup>7</sup>GTP/eIF-(iso)4F system (one short lifetime for ant-m<sup>7</sup>GTP and a longer lifetime for the protein complex). Similar results were obtained for the cap analog/eIF-4F/PABP complex. This suggested that the microenvironment of the cap analogue was similar in the cap analog/eIF and the cap analog/eIF/PABP complexes.

A typical binding isotherm can be developed as described elsewhere (29). For a simple binary reaction:



$$F = \frac{KF_1[\text{eIF}] + F_0}{1 + K[\text{eIF}]} \quad (5)$$

$$[\text{eIF}] = \frac{-(1 + K[\text{cap}^*]_T - K[\text{eIF}]_T) + \sqrt{(1 + K[\text{cap}^*]_T - K[\text{eIF}]_T)^2 + 4K[\text{eIF}]_T/2K}}{2K}$$

Equation 5 can be solved for association constant ( $K$ ) and where cap<sup>\*</sup>, eIF, and cap<sup>\*</sup>·eIF are free fluorescent ant-m<sup>7</sup>-GTP, cap-binding protein, and ant-m<sup>7</sup>GTP/eIF complex, respectively. [cap<sup>\*</sup>]<sub>T</sub> and [eIF]<sub>T</sub> are the total concentrations of ant-m<sup>7</sup>GTP and protein, respectively.  $F$  is the measured value of the fluorescence intensity at any point in the titration;  $F_0$  and  $F_1$  are the fluorescence intensities at the start and at the end of the titration. Thus,  $K$  can be obtained by fitting the experimental data using nonlinear regression analysis. In principle, any one point will give a  $K$  value if the initial and final points are known, there is therefore considerable redundancy in the data fitting.

For the ternary system, the fluorescence data were treated analogously. Binding of eIFs to PABP in the presence of cap analogue can be described as shown in Scheme 1. From our previous data, PABP does not bind to the cap analogue under these conditions.  $K_1$  and  $K_2$  are dissociation constants for ant-m<sup>7</sup>GTP/eIF and eIF/PABP, respectively.  $K_3$  and  $K_4$  are dissociation constants for cap analogue/eIFs to PABP and eIFs/PABP to cap analogue, respectively. Thermodynamically,  $K_1K_3 = K_2K_4$ . In accordance with Scheme 1, the apparent association constant ( $K_{\text{app}}$ ) is derived as follows:

$$\begin{aligned} K_{\text{app}} &= \frac{([\text{cap}^* \cdot \text{eIF}]) + (F_2/F_3)[\text{cap}^* \cdot \text{eIF} \cdot \text{PABP}]}{[\text{cap}^*][\text{eIF}] + [\text{eIF} \cdot \text{PABP}]} \\ &= \frac{([\text{cap}^* \cdot \text{eIF}]) + (F_2/F_3)[\text{cap}^* \cdot \text{eIF} \cdot \text{PABP}]}{[\text{cap}^*][\text{eIF}]} \\ &= \frac{([\text{cap}^*][\text{eIF}] + [\text{eIF} \cdot \text{PABP}])}{[\text{cap}^*][\text{eIF}]} \\ &= \frac{(1/K_1) + (F_2/F_3)(1/K_1)(1/K_3)[\text{PABP}]}{1 + (1/K_2)[\text{PABP}]} \quad (6) \end{aligned}$$

where [PABP] was the concentration of free PABP and  $F_2$

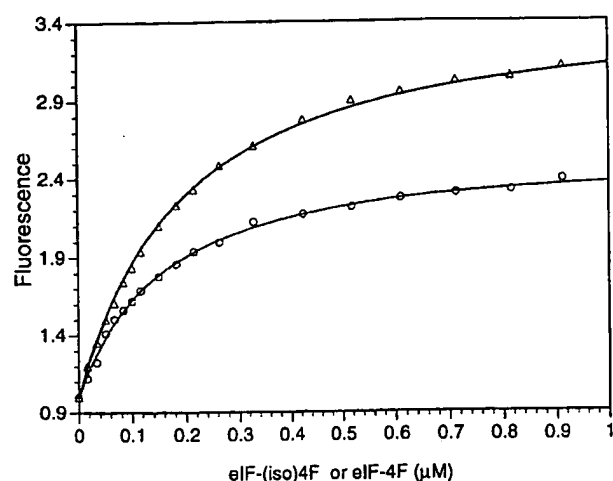


FIGURE 7: Binding of eIF-(iso)4F (triangles) or eIF-4F (circles) to PABP and ant-m<sup>7</sup>GTP. Concentrations of ant-m<sup>7</sup>GTP and PABP were 25 nM and 20 μM, respectively. The solid line represents the best fit to eq 5. Excitation was at 332 nm, and fluorescence emission was integrated from 400 to 500 nm.

and  $F_3$  are fluorescence intensities (integrated from 400 to 500 nm) of ant-m<sup>7</sup>GTP/eIF and ant-m<sup>7</sup>GTP/eIF/PABP complexes, respectively. Under conditions where  $[PABP] \gg [cap^*]$  and  $[PABP] \gg [eIFs]$ ,  $[PABP] = [PABP]_T$ , where  $[PABP]_T$  was the total concentration of PABP. This equation has only one variable,  $K_3$  or  $K_4$ . Titration curves where 25 nM ant-m<sup>7</sup>GTP and 20 μM PABP were titrated with eIF-(iso)4F or eIF-4F are shown in Figure 7. The solid line represents the best fit of the data for eq 5. The  $K_{app}$  value obtained was  $(5.02 \pm 0.12) \times 10^6 M^{-1}$  and  $(8.13 \pm 0.24) \times 10^6 M^{-1}$  for ternary cap ant-m<sup>7</sup>GTP/eIF-(iso)4F/PABP and ant-m<sup>7</sup>GTP/eIF-4F/PABP system, respectively. Theoretically, from eq 6,  $K_{app}$  is a function of  $[PABP]$ . By changing the concentration of PABP, we could solve for all of the  $K_d$  values in Scheme 1 by nonlinear regression analysis without individual binding measurements. However, limitations on the amount of protein needed made this impractical. The fitted values of  $K_3$  for PABP dissociation from ant-m<sup>7</sup>GTP/eIF/PABP complexes were in the nanomolar range, while values of  $K_4$  for ant-m<sup>7</sup>GTP dissociation from the ant-m<sup>7</sup>GTP/eIF/PABP complexes were in the micromolar range. Increasing the concentration of PABP 2-fold did not alter  $K_{app}$  significantly (data not shown). This result was reasonable considering the  $K_1$  and  $K_2$  values obtained from separate titrations since any concentration of PABP used in the micromolar range will be canceled out in eq 6. The exact values of  $K_d$  for eIF-(iso)4F and eIF-4F are summarized in Table 1. The presence of cap analogue enhances the binding eIF to PABP approximately 40-fold. The binding affinities of the eIF/PABP complexes to cap analogue were enhanced approximately 40-fold as compared to the affinity of cap-binding proteins alone. The two cap-binding proteins, eIF-4F and eIF-(iso)4F behaved very similarly. In both cases, the presence of PABP enhanced cap-binding affinity and the presence of the cap-enhanced protein complex formation.

## DISCUSSION

PABP has known RNA binding motifs, but here we have demonstrated protein-protein interactions as well as the

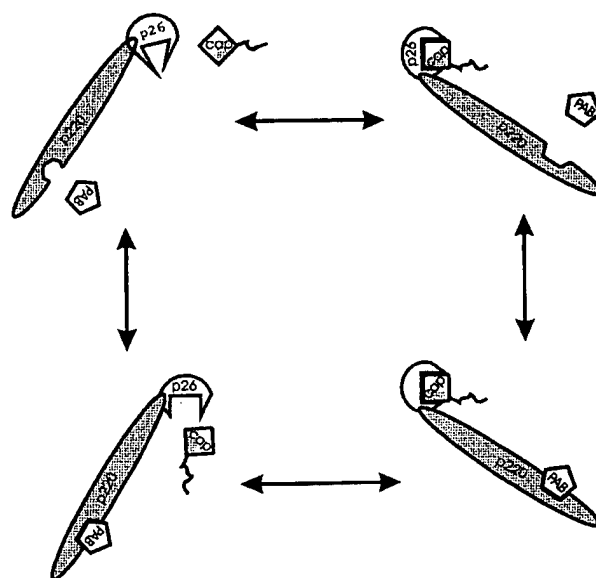


FIGURE 8: Schematic diagram for the interaction of ant-m<sup>7</sup>GTP, eIF-4F, and PABP. The eIF-4F is presented as two subunits, eIF-4G and eIF-4E.

effects of these interactions on cap-binding affinity. PABP has been shown to influence translation, but the mechanism is unclear. Direct fluorescence titrations demonstrated that the interaction of PABP with eIF-4F and eIF-(iso)4F increase the cap binding affinity by an order of magnitude.

Fluorescence energy transfer studies on these cap-binding proteins/PABP complex indicate that they have different binding sites for the 5' cap and 3' poly(A) tail. Even though cap analogue binds to eIF-4F or eIF-(iso)4F with a  $K_d$  of 4–9 μM, in the presence of PABP, cap binding is enhanced 40-fold. Similarly, cap bound to cap-binding protein will enhance the binding to PABP. Since cap analogues have been demonstrated by CD analysis to induce a conformational change in eIF-(iso)4F (30), this may explain the higher affinity of cap-eIF for PABP. The data presented here support a model where eIF-4F, PABP, and cap are involved in formation of an RNA complex, which is summarized in Figure 8. The model shows the microenvironment of the cap to be similar in the ant-m<sup>7</sup>GTP/eIF-4F and ant-m<sup>7</sup>GTP/eIF-4F/PABP complexes, since no significant change in fluorescence intensity or lifetime measurements occurred. The binding of cap to eIF-4F results in conformational changes in eIF-4F, which subsequently enhance the binding to PABP. eIF-4F bound to PABP may have a conformational change in order to bring the nonpolar domains together, which would subsequently enhance cap binding.

The fact that PABP enhances the affinity of eIF-4F or eIF-(iso)4F for the cap analogue is an interesting observation in light of fact that PABP interacts with the 4G subunit (8, 9) and not directly with the cap-binding subunit, 4E. These results suggest that a conformational change is propagated through the 4G subunit to the cap binding subunit, 4E. Enhancement of translation by PABP may be at least partially accounted for by a direct effect on the cap affinity. Additional enhancement may occur through a mechanism whereby the 5' and 3' ends of mRNA interact and ribosomes are more readily recycled.

## REFERENCES

1. Jackson, R. J., and Standart, N. (1990) *Cell* 62, 15–24.
2. Munroe, D., and Jacobson A. (1990) *Mol. Cell. Biol.* 10, 3441–3455.
3. Munroe, D., and Jacobson A. (1990) *Gene* 91, 151–158.
4. Bernstein, P., and Ross, J. (1989) *Trends Biochem. Sci.* 14, 373–377.
5. Sonenberg, N. (1988) *Prog. Nucleic Acid Res. Mol. Biol.* 35, 173–207.
6. Rhoads, R. E. (1988) *Trends Biochem. Sci.* 13, 52–56.
7. Gallie, D. R., and Tanguay, R. L. (1994) *J. Biol. Chem.* 269, 17166–17173.
8. Tarun, S. Z., Jr., and Sachs, A. B. (1996) *EMBO J.* 15, 7168–7177.
9. Le, H., Tanguay, R. L., Balasta, M. L., Wei, C.-C., Browning, K. S., Goss, D. J., and Gallie, D. R. (1997) *J. Biol. Chem.* 272, 16247–16255.
10. Ren, J., and Goss, D. J. (1996) *Nucleic Acids Res.* 24, 3629–3634.
11. Yang, J., and Hunt, A. G. (1992) *Plant Physiol.* 98, 1115–1120.
12. Lax, S. R., Lauer, S. J., Browning, K. S., and Ravel, J. M. (1986) *Methods Enzymol.* 118, 109–128.
13. Lax, S. R., Browning, K. S., Maia, D. M., and Ravel, J. M. (1986) *J. Biol. Chem.* 261, 15632–15636.
14. Odom, O. W., Jr., Robbins, D. J., Lynch, J., Dottavio-Martin, D., Kramer, G., and Hardesty, B. (1980) *Biochemistry* 19, 5947–5954.
15. Bradford, M. M. (1976) *Anal. Biochem.* 72, 248–254.
16. Firpo, M. A., Connelly, M. B., Goss, D. J., and Dahlberg, A. E. (1996) *J. Biol. Chem.* 271, 4693–4698.
17. Locke, B. C., MacInnis, J. M., Qian, S.-J., Gordon, J. I., Li, E., Fleming, G. R., and Yang, N.-C. (1992) *Biochemistry* 31, 2376–2383.
18. Freifelder, D. (1976) *Physical Biochemistry: Applications to Biochemistry and Molecular Biology*, pp 440–441, W. H. Freeman, San Francisco.
19. Weber, G., and Teale, F. W. J. (1957) *Trans. Faraday Soc.* 53, 646–655.
20. Goss, D. J., Carberry, S. E., Dever, T. E., Merrick, W. C., and Rhoads, R. E. (1990) *Biochim. Biophys. Acta* 1050, 163–166.
21. Darzynkiewicz, E., Stepinski, J., Ekiel, I., Goyer, C., Sonenberg, N., Temeriusz, A., Jin, Y., Sijuwade, T., Haber, D., and Tahara, S. M. (1989) *Biochemistry* 28, 4771–4778.
22. Darzynkiewicz, E., Stepinski, J., Ekiel, I., Jin, Y., Haber, D., Sijuwade, T., and Tahara, S. M. (1988) *Nucleic Acids Res.* 16, 8953–8962.
23. Darzynkiewicz, E., Ekiel, I., Lassota, P., and Tahara, S. M. (1987) *Biochemistry* 26, 4372–4380.
24. Lakowicz, J. R. (1983) *Principles in Fluorescence Spectroscopy*, Plenum Press, New York.
25. Birmachu, W., and Reed, J. K. (1988) *Photochem. Photobiol.* 47, 675–683.
26. Marcotrigiano, J., Gingras, A.-C., Sonenberg, N., and Burley, S. K. (1997) *Cell* 89, 951–961.
27. Goss, D. J., Woodley, C. L., and Wahba, A. J. (1987) *Biochemistry* 26, 1551–1556.
28. Sachs, A. B., Davis, R. W., and Kornberg, R. D. (1987) *Mol. Cell. Biol.* 7, 3268–3276.
29. Heyduk, T., and Lee, J. C. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 1744–1748.
30. Wang, Y., Sha, M., Ren, W. Y., van Heerden, A., Browning, K. S., and Goss, D. J. (1996) *Biochim. Biophys. Acta* 1297, 207–213.

BI9724570